Performance-monitoring integrated reweighting model of perceptual learning

Performance-monitoring integrated reweighting model of perceptual learning

Grigorios Sotiropoulos[a], Aaron R. Seitz[b], Peggy Seriès[a]

[a]*School of Informatics, University of Edinburgh*
Edinburgh, UK
[b]*Department of Psychology, University of California, Riverside*
Riverside, CA, USA

Abstract

Perceptual learning (PL) has been traditionally thought of as highly specific to stimulus properties, task and retinotopic position. This view is being progressively challenged, with accumulating evidence that learning can generalize (transfer) across various parameters under certain conditions. For example, retinotopic specificity can be diminished when the proportion of easy to hard trials is high, such as when multiple short staircases, instead of a single long one, are used during training. To date, there is a paucity of mechanistic explanations of what conditions affect transfer of learning. Here we present a model based on the popular Integrated Reweighting Theory model of PL but departing from its one-layer architecture by including a novel key feature: dynamic weighting of retinotopic-location-specific vs location-independent representations based on internal performance estimates of these representations. This dynamic weighting is closely related to gating in a mixture-of-experts architecture. Our dynamic performance-monitoring model (DPMM) unifies a variety of psychophysical data on transfer of PL, such as the short-vs-long staircase effect, as well as several findings from the double-training literature. Furthermore, the DPMM makes testable predictions and ultimately helps understand the mechanisms of generalization of PL, with potential applications to vision rehabilitation and enhancement.

Keywords: perceptual learning, computational, model, specificity, transfer, double training, mixture-of-experts

Email addresses: greg.sotiropoulos@gmail.com (Grigorios Sotiropoulos), aseitz@ucr.edu (Aaron R. Seitz), pseries@inf.ed.ac.uk (Peggy Seriès)

1. Introduction

One of the hallmarks of perceptual learning (PL) is its specificity for various aspects of task and stimulus configuration, such as stimulus orientation or curvature (Poggio et al., 1992; Shiu & Pashler, 1992; Ahissar & Hochstein, 1993; Fahle & Morgan, 1996; Fahle, 1997), retinotopic position (Karni & Sagi, 1991; Shiu & Pashler, 1992; Schoups et al., 1995; Sowden et al., 2002) and ocularity (Schoups et al., 1995). Retinotopic specificity has been recently challenged by various experimental approaches. For example, Xiao et al. (2008) employed a novel double-training paradigm, where contrast discrimination was trained, in an interleaved fashion, with an orientation discrimination task at a second location, and found this enabled transfer of learned improvements to the second location. Since then, several studies from this lab have shown that retinotopic specificity can be abolished by various simultaneous or sequential double-training procedures (Zhang et al., 2010; Wang et al., 2012, 2014). However, recent findings by Hung & Seitz (2014) suggest that retinotopic specificity depends on the details of the experimental procedure and, for example, on the proportion of hard (e.g. near-threshold) trials during training. In particular, Hung & Seitz (2014) found that if training sessions consist of multiple short staircases (whereby there is a significant proportion of trials well above threshold until the staircase starts to converge) learning can transfer across retinotopic positions. If instead a single long staircase (which contains a higher proportion of parathreshold trials) is used, then retinotopic specificity was observed, even under double training procedures.

From a physiological perspective, there is disagreement as to when and why retinotopic specificity is observed in PL. Classically, such specificity has been attributed to plasticity within early visual cortex (Karni & Sagi, 1991; Poggio et al., 1992; Crist et al., 1997), where retinotopic information is best preserved. This position has found some experimental support by Schoups et al. (2001), who showed modest sharpening of the tuning curves of V1 neurons following PL, and by Yang & Maunsell (2004), who found a similar tuning curve sharpening in area V4. However, there is mounting evidence of plasticity in higher, non-sensory areas as well. A notable example is the study of Law & Gold (2008), who found that in a motion direction discrimination task, behavioral improvement correlated with plasticity in neurons in the lateral intraparietal cortex (LIP, a decision area) but not in sensory area MT that was also active during the task. Furthermore, a recent PL experiment of Vernier discrimination that also measured event-related potentials (ERP) found evidence that top-down influences mediate the specificity of PL (Zhang et al., 2013).

The topic of specificity, either to retinotopic position or to task/stimulus features, continues to be a hotly debated matter. A path towards resolving this matter is in the development of theoretical models that can provide mechanistic explanations of how learning manifests differently from different training paradigms. The majority of the published models belong to a category called *reweighting models,* in which learning does not manifest within representations of feature space but instead is realized as weight changes in the connections that allow "read out" of activity of representations areas by higher-level decision units. An early example was the Radial Basis Function model of Poggio et al. (1992)that helped explain perceptual learning of hyperacuity, which was followed by a number of more developed models that employed similar principles (Weiss et al., 1993; Sotiropoulos et al., 2011). The model of Dosher & Lu (1998) popularized reweighting models, and Petrov et al. (2005) extended and provided an in-depth analysis of the multichannel Augmented Hebbian Reweighting Model (AHRM). This model, in a series of publications (Petrov et al., 2005, 2006), was able to account for partial transfer of learning to different stimulus orientations and disruption of learning ("switch costs") when the either the orientation of the stimulus or the context changed. The same model was later used to propose that specificity of perceptual learning with respect to stimulus or task parameters depends on task precision (Jeter et al., 2009), which is a measure of task difficulty (a notion suggested by Ahissar & Hochstein, 2004 in their Reverse Hierarchy Theory).

More recently, the AHRM was extended as the Integrated Reweighting Theory (IRT, Dosher et al., 2013) model, to account for retinotopic specificity and transfer. The model consists of location-invariant and location-specific representations that all connect to a single decision unit and was able to reproduce multiple learning patterns in an orientation experiment. Talluri et al. (2015) presented a multi-layer variant of the IRT model of Dosher et al. (2013), with location-invariant representations being explicitly constructed out of location-specific ones. The model weighted location-invariant representations in proportion to the "confidence" (the magnitude of the output) of the decision unit and suggested that this framework can account for Hung & Seitz (2014)'s observations that double training-induced transfer depends on the precision of stimuli used during initial training.

While these models provided useful insight into the possible mechanisms of transfer of learning across retinotopic locations, they have their limitations. The IRT is not able, as is, to account for the recent findings of Hung & Seitz (2014) (the differential effect of long vs short staircases in

double training conditions) and of Wang et al. (2014) (the "piggybacking effect"). The model of Talluri et al. (2015), on the other hand, is not able to handle a diverse set of tasks and stimuli in a unified, normative manner. These observations pointed out a gap in existing models of PL; to bridge that gap and advance Talluri et al. (2015)'s idea of changing the weight of the location-invariant representation system on time scales smaller than that of PL itself, we propose an extension of the IRT that goes beyond a one-layer architecture by incorporating a mechanism of dynamically (on a time scale of a few tens of trials) assigning weights to responses generated from location-specific and location-invariant visual representations. This dynamic weighting is conceptually very similar to the gating operation performed in a class of neural network models introduced more than two decades ago in the machine learning literature – the mixture-of-experts (MoE, Jacobs et al., 1991b; Jordan & Jacobs, 1991; Hampshire & Waibel, 1992; Jacobs, 1997). The similarities of our model to a MoE architecture give it its unique ability to handle recent experimental findings on retinotopic location transfer of learning. In our model, PL occurs in the readout of both location-invariant and location-specific representations and retinotopic transfer of learning is enabled by dynamically weighting the two readouts based on the recent performance of the location-invariant readout. This in turn depends on the experimental protocol used. Thus our model suggests a radical idea: learning occurs everywhere, it just needs to be "unlocked" by the appropriate experimental procedure. To our knowledge, this is the first time that such a model has been used in the PL literature.

2. Methods

The DPMM consists of 3 representation layers; two retinotopic layers (hereafter called "V1"), assumed to be located in diagonally opposite visual quadrants, and a non-retinotopic layer (hereafter called "V4"). Each representation layer consists of an array of units, accepting a raw stimulus image and having Gabor receptive fields tuned to a particular orientation and spatial frequency. All representation units connect to a single output (or "decision") unit, as in the IRT model (Dosher et al., 2013; cf. Talluri et al., 2015); however, the V1 and V4 representations are also read out seperately, as individual "experts".

The main concept behind DPMM is that the high-level decision unit integrates V1 and V4 inputs in a weighted manner whereby the weight of the V4 subsystem increases with its performance. The dependence of the weight of V4 on its recent performance constitutes multiplicative gating, akin to a mixture-of-experts (MoE) architecture (Jacobs et al., 1991b; Jordan & Jacobs, 1991; Hampshire & Waibel, 1992; Jacobs, 1997). A central assumption here is that, in experiments with feedback,

the brain can not only estimate its overall performance in the recent past but can also estimate the performance of V4 in isolation, by comparing the output of the decision unit to the output of a decision unit that receives only V4 input.

2.1. Learning in AHRM

Learning in the AHRM of Petrov et al. (2005) is Hebbian, augmented by feedback when available. The weight update at the synapse between the i-th representation unit and the decision unit implements soft weight bounding (preventing weights from becoming smaller than w_{min} or greater than w_{max}) and is given by:

$$\Delta w_i = (w_i - w_{min})\,[u_i]_- + (w_{max} - w_i)\,[u_i]_+ \quad (1)$$

$$u_i = \eta X_i (o - \bar{o}) \quad (2)$$

Here o is the output of the model given by the dot product of the weight vector and the vector of activations of the representation units (plus a bias term) passed through a compressive nonlinearity G():

$$o = \mathrm{G}(\mathbf{w} \cdot \mathbf{X} - w_b b + \epsilon_d) \quad (3)$$

$$\mathrm{G}(x) = \tanh(\frac{\gamma_o x}{2}) A_{max} \quad (4)$$

o is negative for "left" responses and positive for "right" ones; $\bar{o}$ is the running average of the output; ϵ_d is Gaussian decision noise of zero mean and variance σ_d; A_{max} is the maximum activation of the decision unit; γ_o is the gain of the activation function. When trial-by-trial feedback is available, the feedback value ($F = \pm 1$) is added to the output before the latter is used in Eq. 1 to update the weight of each connection:

$$o = \mathrm{G}(\mathbf{w} \cdot \mathbf{X} - w_b b + w_f F + \epsilon_d) \quad (5)$$

In the IRT of Dosher et al. (2013), $\mathbf{w} \cdot \mathbf{X}$ refers to both V1 and V4 inputs and is thus equal to $\mathbf{w_{V1}} \cdot \mathbf{X_{V1}} + \mathbf{w_{V4}} \cdot \mathbf{X_{V4}}$ (they use the terms location-dependent and location-independent representations).

2.2. DPMM innovations

The purpose of this study is to investigate whether an AHRM-derived MoE-type of network that receives input from early (V1) and later (V4) visual areas is able to account for the characteristics of transfer of learning across retinotopic locations. For the model to be able to simulate

experiments reported in the relevant literature, a number of changes to the design and its publicly available implementation (at http://alexpetrov.com/proj/plearn/) were necessary. Note that the same two types of populations (V1 and V4) and the same connections between these and the decision unit were used for all tasks and stimuli.

2.2.1. Flexibility in specifying the experimental paradigm

In order to compare different methodologies used so far in literature, we gave each task and condition a numerical ID, and the user is instructed to create simple data structures that encode the entire experiment with all its details (such as number of staircases and their parameters, number and type of sessions etc). Specifying an entire psychophysical experiment through a "template" file enables quicker testing of different scenarios and easier replication of published results. The source code for this will be made publicly available so that interested researchers can easily run their own experiments. For the following simulations, we have specified templates for 3 different tasks: pair-of-Gabors Vernier discrimination, Gabor orientation discrimination and pair-of-Gabors contrast discrimination.

Related to the above, the DPMM can handle both binary (e.g. "in a pair-of-Gabors Vernier stimulus, is the top Gabor to the left or to the right of the bottom?") and 2-IFC tasks ("was the stimulus in the first or in the second presentation more leftward?"), again via simple entries in the template file.

2.2.2. Alternative bias calculation

The original AHRM/IRT includes a bias unit that injects a bias term b in the summation of Eq. 5. b was based on a running average of "left/down/clockwise" (-1) or "right/up/counterclockwise" (+1) responses during training. Assuming that left/right, up/down and clockwise/counterclockwise stimuli appear with the same frequency (which is the case in all simulated experiments considered here), the bias term keeps track of any asymmetry in the model responses and is added to Eqs. 3 and 5 to equalize the responses and stabilize the network.

However, we found that merely requiring an equal number of -1 and 1 responses was not enough in our simulations: responses might be symmetric but the magnitudes of the weighted input $B = \mathbf{w} \cdot \mathbf{X} - w_b b$ (prior to the squashing nonlinearity G(), see Eq. 3) may be asymmetric, i.e. unequal for opposite offsets. For example, at a certain time during training, a Vernier offset of -8 arcmin might result in $B = -0.2$ but the opposite offset of 8 arcmin might result in $B = 0.3$. Such asymmetries might occur in a way that the responses remain counterbalanced but the decision unit output o (prior

to thresholding) still develops a bias. In our simulations we found that, beyond a certain point, this bias caused instability and sharply degraded performance. We note that these asymmetries were observed in the combination of stimuli and representations specific to our instantiation of the model and thus may not apply in contexts that the AHRM/IRT has been used in literature, particularly those that have used different stimuli or representation subystems (the AHRM is modular; the representation subsystem and the decision/reweighting subsystem are dissociable).

To avoid instability, equal but opposite stimulus offsets should result in equal but opposite responses. The computation of the new bias term helps exactly with this problem. It is performed in the same way as the original bias term – by exponentially discounting past bottom-up evidence (as opposed to past responses, see Eq. 14 and 15 in Petrov et al., 2005):

$$b(t+1) = \rho B(t) + (1-\rho)b(t) \tag{6}$$

$$B(t) = \mathbf{w}(t) \cdot \mathbf{X}(t) - w_b b(t) \tag{7}$$

The bias weight w_b is set to 1 throughout the simulations, effectively eliminating a model parameter from our simulations (the bias weight is a necessary feature of previous instantiations of the AHRM that can account for certain subtleties in the data, e.g. Petrov et al., 2005). Note that b is computed separately for V1 and V4 as well as for their combined input. We note that by choosing suitable representations and parameters for our stimuli (described in 2.2.4 and 2.4), we have established such symmetry to a high degree already; however the existence of this mechanism acts as a further failsafe, especially for future stimuli that could be tested with the model.

Apart from the stabilization of learning, the new bias mechanism also plays the role of a "sliding average" of output activity, which, as Petrov et al. (2005, p732 footnote) point out, is physiologically grounded (Bienenstock et al., 1982) and, as we discuss in the next paragraph, mitigates the need for an explicit soft weight-bounding rule. We note that this mechanism is very similar to a separate (to the response bias) running average of post-synaptic activity $\bar{o}$ that has been used in some implementations of the AHRM (e.g. Lu et al., 2010) – the only difference being that the running average in our case is of the value prior to the output nonlinearity.

2.2.3. Alternative weight update

The original AHRM weight updates are "soft bounded" from above and below according to Eq. 1: the closer the weight gets to its maximum (or minimum) value, the smaller the further increase (or decrease), so that the limits are approached asymptotically. While this rule has been successfully

employed in previous implementations of the AHRM, and has even been necesary in modeling alternating training environments (Petrov et al., 2005), the present analysis has shown that it is not appropriate for modelling certain combinations of tasks and stimuli, at least with the representation subsystem most often used in the AHRM/IRT literature: channels tuned to 7 orientations and 5 spatial frequencies for a total of 35 channels. With stimuli such as those in Hung & Seitz (2014) as input, the original weight update rule almost entirely fails to increase the connection weights for channels that *are* informative for the task: under the old rule, the network cannot learn the two-grating Vernier or contrast discrimination tasks because *all* weight traces change very little over the course of training.

We thus removed the soft weight-bounding mechanism from the model, replacing Eqs. 1 and 2 with the classical Hebbian update rule

$$\Delta w_i = \eta X_i o \tag{8}$$

This simpler weight update rule made learning of Vernier and contrast discrimination tasks possible (see AppendixA for a more in-depth discussion of the learning rules). In all simulations described here, weights never approached the higher or lower bounds (they reached at most 50% of either bound, i.e. ±0.5).

2.2.4. Stimulus-dependent integration centers

Whereas the changes thus far refer to the interface between representations and decision unit, the change described here involves a property of the representation subsystem itself: the spatial pooling (integration) kernels. Because stimuli in previous literature were typically circular gratings (embedded in larger "oriented noise" gratings), the energy output of each representation unit was computed via a weighted average of the local energies at all points in space (e.g. all 4096 pixels of the filtered 64×64-pixel images). In previous literature, the weighting kernel has been assumed to be a radially symmetric Gaussian centered at the center of the stimulus, e.g. see p742 of Appendix A in Petrov et al., 2005, who note that "*in the interest of parsimony, and consistent with task demands, each phase-invariant map is then pooled across space over a region comparable with the diameter of the experimental stimuli.*" However, in the case of stimuli that involve more than one components (such as a two-Gabor Vernier), it is reasonable to assume that the integration kernel should be multimodal, for example having one peak for each grating component. Petrov et al. (2005) themselves mention a similar approach, although for a different reason:

"*The spatial pooling reduces the number of representational units $A(\theta, f)$ to 35. Although the*

stimuli in the experiment are presented either above or below the fixation point, this positional uncertainty is not implemented in the model. An extended version with two independent pools of representational units with receptive fields above and below the horizontal meridian would duplicate the present model."

In the DPMM, instead of using two independent pools of representational units for the two Gabor components of the Vernier stimuli, we use one pool and a Gaussian mixture integration kernel (with equal mixing coefficients) with two centers near two diagonal corners of the (filtered) image. Figure B.10, middle shows the Gaussian kernel used in literature so far; the left panel shows the mixture kernel used for the Vernier task modeled here. The kernel centers in this case are diagonally positioned in order to provide a graded, monotonic response to stimulus offsets, ranging from the most negative to the most positive. For example, consider a vertical Vernier with a high positive offset (top grating is to the right of bottom). Because the Gabor components are close to the integration centers, the energy of that representation unit (pooled across space) will be high (say 5). As offset decreases, the gratings move away from the integration centers and toward the vertical central line, at zero offset. The energy in that case will have an intermediate value (say 3). As offset increases in the opposite direction, becoming negative, the Gabors move even further away from the centers and the energy decreases even more (say to 1).

By choosing the integration kernels appropriately, the network can be made to produce a signed output that is a monotonic function of the (signed) stimulus value, and thus offers a natural mapping from stimulus offset to output magnitude. The kernel used for the Gabor orientation task was the same as in previous literature (radially Gaussian at the center of the stimulus). In the case of the contrast stimulus, each blob is the same Gaussian kernel as for the orientation-task stimulus. The spatial configuration of these kernels allows a monotonic stimulus-response mapping, from one extreme ("bottom grating of the highest contrast") to the other ("top grating of the highest contrast"). AppendixB provides more details on the integration kernels and how they allow for monotonic, symmetric network responses.

2.2.5. Output activation function

Both in the orientation task in the model of Petrov et al. (2005) and in all 3 of our tasks, the bottom-up input to the decision unit $\mathbf{w} \cdot \mathbf{X}$ is linear in offset/orientation difference to an excellent approximation and for a wide range of parameter values of the representation subsystem (this is also the case in the model of Weiss et al., 1993). The output activation function (Eq. 4) that has

been used in the AHRM transforms this straight line into a sigmoid ranging from -1 to 1 with the inflection point at zero. In the positive and in the negative range of offsets, the curve is convex and concave, respectively. There are two problems with applying this activation function to our tasks, described in detail in AppendixC. Briefly, usage of the current function in the way we intend for our tasks results in model psychometric functions that do not match well the empirical ones and, furthermore, it makes the model sensitive to staircase parameters.

To address these problems we replaced Eq. 4 with a function (Eq. 10) that has a sigmoid shape in both the negative and the positive parts of the output and shows saturation effect both near zero and near large offsets. Furthermore, noise is now added after the nonlinearity is applied, and the output is constrained in the $[-1, 1]$ range. The new output function results in a realistic psychometric function across all simulated tasks presented here. The parameters γ and c control the slope and location of the sigmoid, respectively.

The same issues apply to the representation activation function, which is identical to the output activation function; in fact, the second problem (sensitivity to staircase details) is more severe due to noise being added before the sigmoid activation function in the representation units than in the decision unit. The reason is that, for the model to exhibit asymptotic performance, a certain amount of representation noise is necessary: as Petrov et al. (2005) point out, if all internal noise is just decision noise, the optimal weights can always increase to overcome it and thresholds can get arbitrarily low. Petrov et al. (2005) deal with the problem by using a soft-bounding weight update rule; as we explained in 2.2.3, we had to change that rule in favor of simple Hebbian updating. Therefore representation noise, which scales with the weights and thus can never be overcome, is a necessary feature in the model. We thus also used the new activation in the representation units, with the same slope (gain) parameter value (thus saving one model parameter) and a slightly smaller location parameter. Figures E.12 and E.13 of AppendixE show activities of the representations and the decision unit in the contrast discrimination task before and after the nonlinearity is applied.

$$\mathrm{G}_0(x) = \begin{cases} \frac{1}{1+e^{-\gamma(x+c)}} - \frac{1}{1+e^{-\gamma c}} + \epsilon_d, & \text{if } x < 0 \\ \frac{1}{1+e^{-\gamma(x-c)}} - \frac{1}{1+e^{\gamma c}} + \epsilon_d, & \text{otherwise} \end{cases} \tag{9}$$

$$\mathrm{G}(x) = \min(\max(\mathrm{G}_0(x), -1), 1) A_{max} \tag{10}$$

It should be noted that this different output activation function has not been necessary in any prior AHRM study because the stimulus property that was varied there was not the orientation

difference (in which the model response is linear) but rather the stimulus contrast. It may well be the case that the psychometric function of model performance with respect to contrast is indeed a sigmoid in these studies, regardless of the shape of the activation function.

2.2.6. V1/V4 weighting through V4 performance monitoring

The changes to the IRT presented thus far were incremental and specific, contributing mainly to its ability to handle our various tasks and stimuli and produce realistic psychometric functions. The final change, however, is a radical one and central to the model's ability to predict in what conditions retinotopic location transfer should occur. It is also the feature that makes DPMM depart from the one-layer feed-forward architecture of the IRT.

As in Talluri et al. (2015), we introduce weights on the output of V1 ($\mathbf{w_{V1}} \cdot \mathbf{X_{V1}}$) and V4 ($\mathbf{w_{V4}} \cdot \mathbf{X_{V4}}$) subsystems before summing those outputs (these weights are not to be confused with the synaptic weights $\mathbf{w_{V1}}$ and $\mathbf{w_{V4}}$). The model, apart from computing the weighted sum of V1 and V4 outputs, also computes the output that V4 on its own would have and maintains a running average of V4 performance, based on trial-by-trial feedback, in a manner identical to the running average of responses for the bias calculations (Eqs 14 and 15 in Petrov et al., 2005). In particular, the running average of V4 performance in trial $t + 1$ is given by

$$p_{V4}(t+1) = \begin{cases} (1-\rho_{V4})p_{V4}(t) + \rho_{V4}, & \text{if } \mathrm{sgn}(\mathrm{G}(\mathbf{w_{V4}} \cdot \mathbf{X_{V4}} - b_{V4})) = \mathrm{sgn}(F) \\ (1-\rho_{V4})p_{V4}(t), & \text{otherwise} \end{cases} \tag{11}$$

Here ρ_{V4} is a rate parameter, analogous to ρ in Eq. 6. The higher the performance of V4, the larger the weight given to V4 in the summed output. This mechanism avoids any ad-hoc mapping between stimulus difficulty and V4 weight, instead looking only at the performance of V4 over a number of recent trials (with past trials exponentially discounted). The weight assigned to V4 in the decision process in each trial is given by a sigmoid function that ranges from 0 to 1 (Figure 1):

$$k_{V4}(t) = \frac{1}{1 + \exp(-\beta(p_{V4}(t) - p_{HH}))} \tag{12}$$

$$k_{V1}(t) = 1 - k_{V4}(t) \tag{13}$$

where p_{V4} is the running average of V4 performance; β is the slope of the sigmoid, determining how quickly k_{V4} changes with p_{V4}; where p_{HH} is the center (or "location") of the sigmoid, i.e. the performance at which V1 and V4 are assigned equal weights ($k_{V1} = k_{V4} = 0.5$). Thus, in each trial,

the output of the decision unit is given by

$$o = \mathrm{G}(k_{\mathrm{V1}}\mathbf{w_{V1}} \cdot \mathbf{X_{V1}} + k_{\mathrm{V4}}\mathbf{w_{V4}} \cdot \mathbf{X_{V4}} - b + w_f F) \tag{14}$$

This equation shows the fundamental difference between the DPMM and the AHRM/IRT, which is the dynamic multiplicative weighting of V1 and V4 implicit in the term $k_{\mathrm{V4}}\mathbf{w_{V4}} \cdot \mathbf{X_{V4}}$. This multiplicative weighting is conceptually similar to a mixture-of-experts (MoE) architecture (Jacobs et al., 1991b; Jordan & Jacobs, 1991; Jacobs & Jordan, 1993; Jacobs, 1997), where the experts are "gated" by a separate gating network that receives the same inputs as the experts and acts as a "stochastic switch" that chooses the i-th expert with a probability that is a function of the inputs connected to the i-th gating output. Both the experts and the gating network are trained with gradient descent, which produces an indirect coupling between experts and the gating network (Jacobs et al., 1991b). A comparison between the DPMM and mixture-of-experts models is made in the Discussion.

The location of the sigmoid p_{HH} is a measure of the tradeoff between task performance and generalizability of learning (presumably a desirable property of the brain). If performance was the only objective, the brain would always resort to the optimal mixture of V1 and V4 (giving more weight to V1, in proportion to its relative precision, i.e. the ratio of output variances of V4/V1) in order to deal with all but the easiest tasks. Allowing p_{HH} to vary can potentially explain the interindividual variability in location transfer. For example, "generalizers" might unconsciously consider more important the applicability of their acquired skill in novel situations, whereas "specializers" might seek maximum performance in a particular situation. Note that p_{V4} and k_{V4} are only updated during training sessions, where trial-by-trial feedback is given; in test sessions, the values of these components are the same as those at the end of the immediately preceding training session. This mechanism could also be used in experiments that provide block feedback, in which case p_{V4} would be readily available, or even when no feedback is present, as long as the task is simple enough so that an internal performance estimate can be computed.

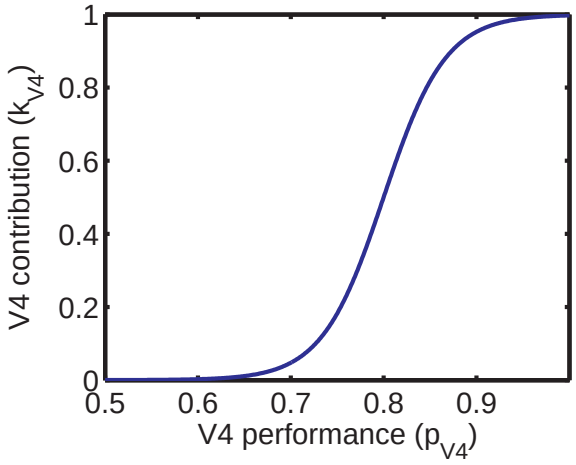

Figure 1: Contribution (aka weight) of V4 input in the decision unit as a function of V4 performance (Eq. 12).

Values of the parameters in Eq. 12 that produce good fits to all data sets examined in the present work are $\beta = 30$ and $p_{HH} = 0.8$. In cases where there is location transfer in the single-staircase condition, the best-fit value of p_{HH} would be lower (e.g. 0.77) because in that condition p_{V4} is expected to be less than with multiple staircases.

2.3. Validation of model changes

To ensure that the aforementioned changes in the model did not diminish its explanatory power for previous datasets, we fit the DPMM to the data of Dosher et al. (2013). In this study, observers were asked to perform fine orientation discrimination (±5°) of Gabor patches around a particular reference orientation and retinotopic location for 8 sessions ("Training" phase). In the next 8 sessions ("Transfer" phase), observers were asked to perform the same task either at a different reference orientation (group "Switch O"), at a different position (diagonal retinotopic location; group "Switch P") or both (group "Switch OP"). The purpose of these switches was to measure the extent of transfer of learning across stimulus orientations (where specificity of learning is the norm), retinotopic locations and both. Each group performed the task with noise-free stimuli as well as stimuli degraded by Gaussian noise. Using staircase procedures, the authors measured contrast thresholds, i.e. the stimulus contrast at which observers could respond correctly in 80% or 70% (depending on the staircase rule used) of the trials. As in Dosher et al. (2013), a single set of parameters was used for both the noise-free and the noisy stimuli, as each observer group was exposed to both, in an intermixed fashion. Whenever possible, we kept the same parameter values as those reported by Dosher et al. (2013). It turned out that most parameters did not need to be changed; out of the 24 parameters reported in Table 1 of Dosher et al. (2013), only the 6 parameters "Adjusted

for the data" and the scale of the initial weight vector were adjusted in the DPMM. Importantly, the tuning parameters (orientation and spatial frequency bandwidths and spacing) of the representation system were kept the same (see AppendixD for more details regarding model parameters).

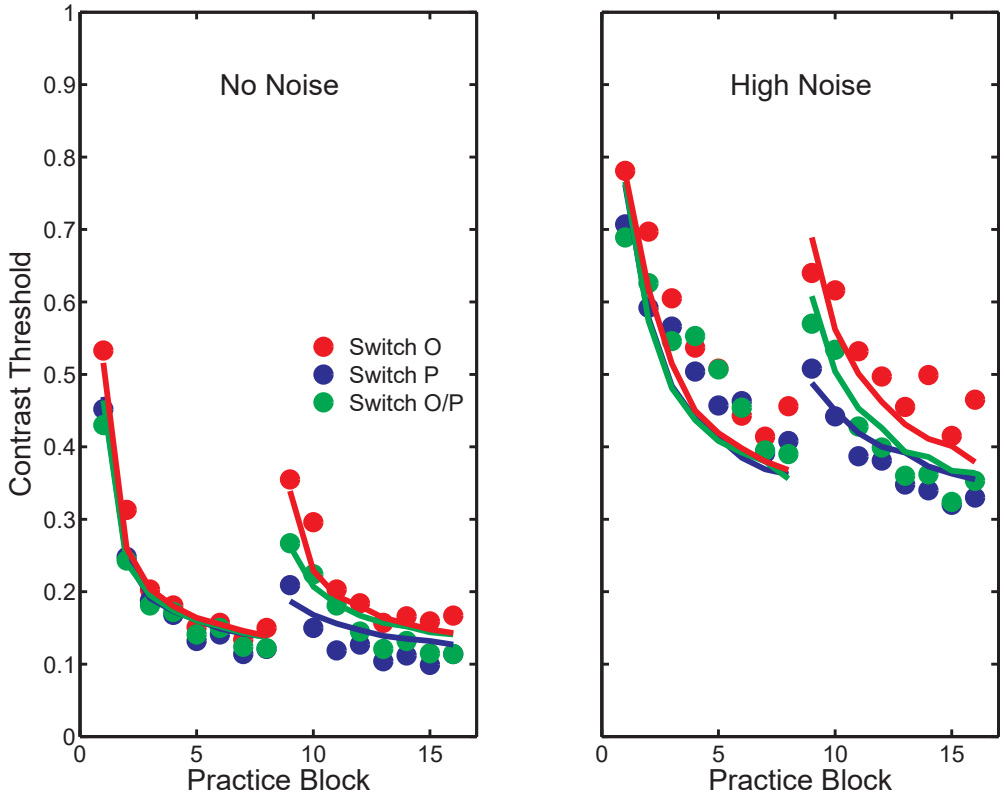

Figure 2: Contrast thresholds for the initial training phase (blocks 1-8) and subsequent practice during the transfer phase (blocks 9-16) in an orientation discrimination task after the task switch of the position (P), orientation (O), or both (OP). Dots are data from Dosher et al. (2013); lines are DPMM fits. Left and right panel show results for stimuli (Gabor patches) corrupted by zero or high (external) noise, respectively.

Figure 2 shows the resulting fits (lines) to the psychophysical data of Dosher et al. (2013) (filled circles) and is directly comparable to Figure 2 of that study. The above results are averages of 400 simulation runs per group. Note that we did not perform an exhaustive search of the parameter space, as we were not after the best possible quantitative fit. Our purpose was to show that changes to the model do not impair its predictive power, which has been demonstrated in previous studies. Indeed, the model reproduced the main feature of the Dosher et al. (2013) dataset, namely the different switch costs of the three groups at the beginning of the Transfer phase (block 9), at least as well as the original model. As in Dosher et al. (2013), the qualitative pattern of the data (ordinal pattern of switch costs) can be reproduced under a wide range of parameter values.

The fact that our changed model fits the Dosher et al. (2013) dataset, and did so in a straight-

forward manner, is evidence that we have at least preserved the essence of the AHRM/IRT.

2.4. Model parameters

2.4.1. Tuning and performance differences between V1 and V4

The assumption underlying the mechanism described in 2.2.6 is that V1 shows tuned responses to trials with more subtle stimulus differences than V4 can, as it is assumed to be more precise and less noisy. How does noise and the various tuning properties of the units in V1/V4 affect performance in that area? For noise, the higher the std. dev. of the (additive Gaussian) noise in V1 and V4 units, the worse this layer will perform. How tuning properties affect performance however is less obvious and therefore we tested them seperately: we set all representation parameters (such as orientation domain and bandwidth, spatial frequency domain and bandwidth) but one to the same value, and then measured via simulations performance in each layer. Performance was evaluated by running a simulation of an entire experiment and recording the proportion of correct responses of each layer in all trials, averaged over hundreds of simulations. The results were interesting, among other things because they were stimulus/task dependent. Table 1 shows details. Increasing orientation selectivity only helps with the orientation task and degrades performance in the Vernier task. This is due to the fact that, with higher orientation selectivity, fewer units are activated to a significant degree by the Vernier stimulus, compared to the orientation stimulus. Another interesting finding is that broad tuning for s.f. is better than fine tuning, and this is true for all tasks. The property with the highest impact however was the s.f. domain: changing the 5 spatial frequencies from {1, 1.4, 2, 2.8, 4} cpd to {2, 2.8, 4, 5.7, 8} cpd caused the highest improvement in performance in all tasks (approximately 6%).

	Vernier perf	Ori perf	Contrast perf
↑orientation bandwidth	↑	↓	↑
↑spatial frequency domain	↑	↑	↑
↑spatial frequency bandwidth	↑	↑	↑
↑radial kernel width	↓	↓	↓
↑scaling factor	↑	↑	↑
↑representation noise	↓	↓	↓

Table 1: Effect of changing a representation property (while keeping all others the same) on performance of a layer (V1 or V4) on various tasks (Vernier, orientation and contrast discrimination). For a brief explanation of these parameters, see Dosher et al. (2013); for more details, see Petrov et al. (2005).

Taking all this into account, and in order to maintain a performance difference between V1 and V4 that is roughly independent of task/stimulus, the representation properties were chosen as

shown in Table 2. The choice was made *a priori*, i.e. it was not fit to the data of Hung & Seitz (2014) – it was taken from literature (e.g. Dosher et al., 2013) either verbatim or slightly modified (but within published, physiologically plausible ranges). Essentially, V1 units are more sharply tuned to orientation, less noisy and spatially more precise with smaller receptive fields than those in V4 (receptive field size is inversely proportional to spatial frequency in the AHRM, Petrov et al., 2005). Apart from the orientation[1] and spatial frequency parameters, the rest of the parameters in Table 2 were set *a priori* so that the network output is in a range such that the largest offsets (or orientation/contrast difference) used in the simulation (e.g. ±19 arcmin for Vernier, which are 1.5-3 times above typical thresholds in humans) saturate the decision unit. This allows a wide range of suprathreshold trials to be handled successfully by the model. For example, any Vernier offset above 12 arcmin produced high enough output activation to almost always give the correct response. This is important because it makes the model more robust to the details of the staircases, such as maximum number of trials and reversals, used in the various experiments simulated hereafter (see 2.2.5 for details).

Parameter	Name in source code	Value
Orientation domain θ (degrees)	D_orient	$\{-90, -75, ..., -15, 0\}$
V1 orientation half-width $h_{\theta V1}$ (degrees)	HW_orient	15
V4 orientation half-width $h_{\theta V4}$ (degrees)	HW_orient	24
V1 spatial frequency domain f_{V1} (cpd)	D_sqfreq	{2, 2.8, 4, 5.7, 8}
V4 spatial frequency domain f_{V4} (cpd)	D_sqfreq	{1, 1.4, 2, 2.8, 4}
V1 spatial frequency bandwidth h_{fV1} (octaves)	octaves	1
V4 spatial frequency bandwidth h_{fV4} (octaves)	octaves	1.6
Vernier radial kernel width h_r (dva)	FW_integr	0.7
Contrast radial kernel width h_r (dva)	FW_integr	1.5
Orientation radial kernel width h_r (dva)	FW_integr	2
Semisaturation constant s^2	CGC_const	0
Representation activation gain γ_r	rep_gain	10
Representation activation location c_r	rep_loc	0.5
Representation scaling factor a	scale_factor	0.1
Maximum activation A_{max}	max_act	1

Table 2: Representation layer parameters set *a priori* and used in all simulations. A brief explanation of these parameters is given by Dosher et al. (2013); for a more in-depth description, see Petrov et al. (2005). Parameter values are the same for V1 and V4 except when otherwise specified.

[1]In the simulation of the orientation task of Hung & Seitz (2014) the reference orientation fell in the middle ($-45°$) of the orientation domain and equidistantly from the vertical and horizontal gratings in the Vernier conditions (which were at $0°$ and $-90°$ in the sequential training experiment of Hung & Seitz, 2014), under the convention that vertical orientation is $0°$ and clockwise is positive. Using the experimental reference orientation ($-36°$) would result in an asymmetry in the model for no useful purpose.

2.4.2. Initial weights

An important question in reweighting models is how to account for initial performance, which is typically above chance as soon as the feature to be discriminated/detected becomes easy enough. Following previous AHRM literature (Petrov et al., 2005; Dosher et al., 2013), we initialized the weights in a simple, general way to account for baseline performance. In particular, the weight for the representation unit with preferred orientation θ_i is

$$w_i = \mathrm{sgn}(\theta_i + 45^\circ) w_{init} \tag{15}$$

The weights for the 7 orientations from $-90^\circ$to 0° were thus $(-1, -1, -1, 0, 1, 1, 1) \times w_{init}$ with zero weight at the unit with preferred orientation -45° and this vector is replicated for all 5 spatial frequencies. w_{init} is fit to the data and reported in Table 4.

2.4.3. Other model parameters

The following two tables list the rest of the model parameters used hereafter, except when mentioned otherwise. Some of these parameters were set *a priori* according to considerations outlined in 2.4.1 and are listed in Table 3. The parameters that were optimized for the dataset we wished to reproduce are listed in the next section (Table 4).

Parameter	Name in source code	Value
Running average rate ρ	runav_rate	0.02
V4 performance running average rate	V4runav_rate	0.0024
V4 performance sigmoid location p_{HH}	perf_HH	0.8
V4 performance sigmoid slope β	perf_slope	30
Feedback weight w_f	fdbk_wgt	0.4
Feedback fraction	fdbk_frac	1
Feedback mode	fdbk_mode	3
Output activation function gain (slope) γ_o	out_gain	10
Maximum activation A_{max}	max_act	1

Table 3: Parameters of the IRT model set *a priori* in all simulations. For a brief explanation of these parameters, see Dosher et al. (2013); for more details, see Petrov et al. (2005).

2.4.4. Fitted parameters

All model fitting, which involved 7 free parameters, was performed on the first 7 sessions of the sequential training data of Hung & Seitz (2014, Figure 4) using least squares. These sessions are essentially instances of single training, i.e. they serve as control sessions meant to isolate the effect of staircase mode. To fit this relatively large number of free parameters, we used nested grid search

constrained by empirical findings. We then performed a large number ($n = 1000$) of simulation runs and computed average thresholds.

3. Results

3.1. Simulations of previous experiments

3.1.1. Single vs multiple staircases in sequential training: IRT model

We first assessed whether a model whose output is a nonlinear function of the sum of dot products $\mathbf{w_{V1}} \cdot \mathbf{X_{V1}} + \mathbf{w_{V4}} \cdot \mathbf{X_{V4}}$ (see Eq. 5), as is the case in the original IRT model of Dosher et al. (2013). The model can account for the differential transfer of learning in single versus multiple-staircase conditions found by Hung & Seitz (2014). We thus disabled the new weighting mechanism in the DPMM and fixed the V1 and V4 weighting coefficients (k_{V1} and k_{V4}) to 0.5 each before fitting the 7 free parameters (Table 4). Note that the weights $\mathbf{w_{V1}}$ and $\mathbf{w_{V4}}$ are allowed to evolve naturally during training, as in the IRT.

Parameter	Name in source code	Value
V1 representation noise std. dev. σ_{V1}	rep_noise	0.22
V4 representation noise std. dev. σ_{V4}	rep_noise	0.23
Decision noise std. dev. σ_d	out_noise	0.1
V1 initial weight scaling w_{initV1}	W_init	0.17
V4 initial weight scaling w_{initV4}	W_init	0.15
Learning rate λ	learn_rate	8×10^{-5}
Output activation function location c_o	out_loc	0.6

Table 4: Parameters of the IRT model fitted to the sequential training data of Hung & Seitz (2014).

In the sequential double training experiment of Hung & Seitz (2014), where subjects are trained for 5 sessions on a Vernier of one orientation, the IRT model can account for the significant transfer of learning across retinotopic locations that is seen in session 7 (mid-test) following training with Vernier of one orientation and preceding training on with an orthogonal Vernier of orthogonal orientation, in both the multiple and single staircase condition. (Figure 3). In the latter case, learning in the transfer condition is quantitatively similar to that in the data. This, however, was achieved at the cost of replicating the transfer index (TI, the mean % improvement, or MPI, at the transfer location divided by the MPI at the trained location). A $TI \geq 1$ denotes complete transfer whereas a $TI \leq 0$ denotes complete specificity (absence of transfer). TIs for the multiple and single staircase simulations are virtually identical: 0.58 and 0.53, respectively. These numbers quantify the qualitative pattern of the IRT model results, namely that the degree of transfer of learning

does not depend on the staircase condition. This pattern is in conflict with the data of Hung & Seitz (2014), in which *TI*s were significantly modulated by staircase condition (1.19 and −0.12, respectively).

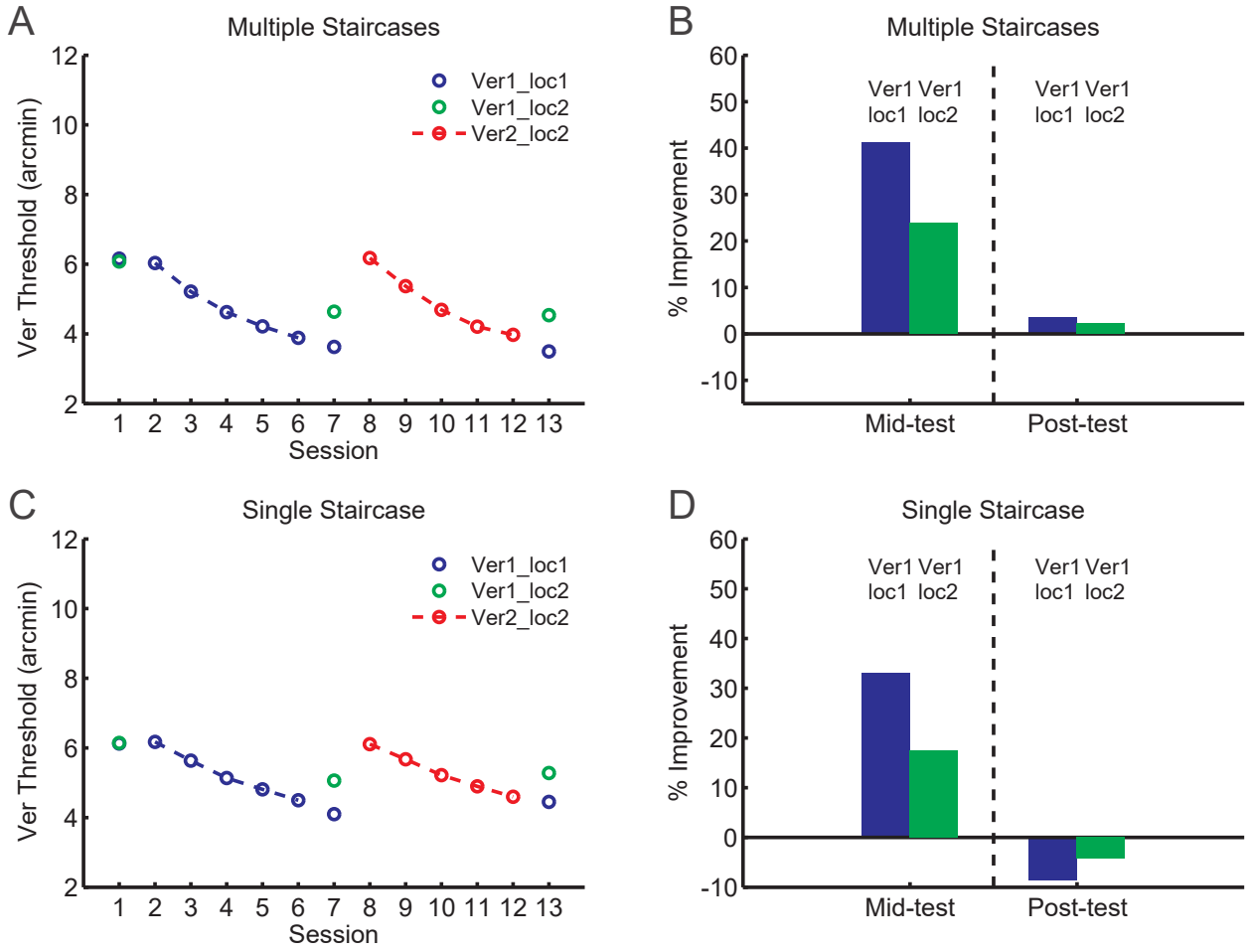

Figure 3: IRT model: results of simulations examining whether learning of a Vernier task of one orientation (Ver1_loc1) followed by learning of a Vernier task of an orthogonal orientation at another retinotopic location (Ver2_loc2) transfers to a Vernier task of the former orientation at the latter location (Ver1_loc2). Thresholds at the end of the 3 test (1, 7, 13) and 10 (2-6, 8-12) training sessions (A and C), as well as MPIs (B and D), are shown for the multiple and single staircase conditions, respectively. Error bars are omitted as they are smaller than the symbols themselves. Cf. Figure 4 of Hung & Seitz (2014) (reproduced in Figure E.14 of AppendixE).

3.1.2. Single vs multiple staircases in sequential training: DPMM

We next examined the impact of re-enabling the dynamic V1/V4 weighting mechanism. To accomplish this, we readjusted 2 of the 7 free parametes previously fitted with the IRT model – the representation noise standard deviations for V1 and V4 (Table 5); these parameters were re-fit because the differential weighting of V1 and V4 changes the overall performance of the system. For example, when the contribution of the V4 component is high, the overall performance would be lower than in the IRT model, if all else is equal. The lower values for the representation noise in

the DPMM reflect the fact that V4 was often given more weight than V1. The other 5 parameters were unchanged, which serves to show that the model is not very sensitive to choices of parameter values.

Like the confidence-based IRT model (Talluri et al., 2015), the DPMM is able to provide a good account of the main results of Hung & Seitz (2014, Figure 4), surpassing the IRT model ($TI = 0.9$ and 0.46 for the multiple and single staircase condition, respectively). The qualitative pattern of results – namely that multiple staircases result in higher contribution from V4, and therefore much greater location transfer, compared to a single staircase – is present under a wide range of parameter values. Suboptimal choice of parameters can diminish the difference between single and multiple staircases, to the point of insignificance; however, no parameter choice can reverse the pattern. All experimental parameters (stimulus properties, procedure and analysis) are identical to those in Hung & Seitz (2014).

Parameter	Name in source code	Value
V1 representation noise std. dev. σ_{V1}	rep_noise	0.14
V4 representation noise std. dev. σ_{V4}	rep_noise	0.22
Decision noise std. dev. σ_d	out_noise	0.1
V1 initial weigth scaling w_{initV1}	W_init	0.17
V4 initial weigth scaling w_{initV4}	W_init	0.15
Learning rate λ	learn_rate	10^{-4}
Output activation function location c_o	out_loc	0.6

Table 5: Parameters of the DPMM fitted to the sequential training data of Hung & Seitz (2014).

Despite the imperfect match of model and data (cf. Figure 4 and Figure E.14 in AppendixE), there are a number of interesting observations. In the multiple staircase condition (panels A and B), learning of Vernier of one orientation (Ver1_loc1) transferred to another retinotopic location (Ver1_loc2) but not to the orthogonal orientation (Ver2_loc1): the latter started off at a high threshold in session 8, just like the Ver1 at the beginning of the experiment. However, by the end of training, the network was able to learn both tasks (two orientations) simultaneously in the second location, as seen in the data (Hung & Seitz, 2014, Figure 5). This is because different representation units are active with Ver1, compared to Ver2 (in both areas V1 and V4), and thus the directions of the optimal weight vectors for these tasks are different. We note that interference is not zero; this is evident in the thresholds at post-test, which are slightly higher than at mid-test. This is due to the fact that Ver1 increases the weights of units that in Ver2 are essentially unresponsive to the stimulus and thus contribute only noise. This is a robust effect in the model but is not prominent in

the data of Hung & Seitz (2014), in which Ver2 in single-staircase mode starts off at a low threshold. However, the multiple-staircase mode data does suggest a switch cost (magenta circles, Figure 4D, Hung & Seitz, 2014). In this sense, there is qualitative agreement between model and data, as the switch cost in the model is higher in the multiple staircase condition. Model results are also comparable to the data of the sequential training experiment of Dosher et al. (2013) simulated in 2.3.

In the single staircase condition, there is no location transfer of Vernier learning in the last test session (post-test), suggesting that the training of V1 with Ver2 at the second location did not help with Ver1 at that location. This is because Ver1 affects a different subset of weights than does Ver2, which is exactly what allows concurrent learning of the two conditions. In multiple-staircase mode, the post-test Ver1 threshold at the second location is actually slightly higher than the mid-test threshold because the weights optimized for Ver2 at the second location only increase noise for Ver1. This is only seen in multiple staircases because V4 is more noisy than V1 (the V4 coefficient at the end of multiple-staircase training was $k_{V4} \approx 0.82$ whereas at the end of single-staircase training $k_{V4} \approx 0.02$).

If we were after a quantitative fit for all data subsets in the sequential training experiment of the Hung & Seitz (2014) study, the parameters of the V4 contribution sigmoid (Eq 12) could be adjusted so that the model fits the "Transfer" subgroup of the single-staircase group – that is, the subgroup that showed complete transfer at mid-test (Figure 5C of Hung & Seitz, 2014). What would not be straightforward to fit, however, is the data in Figure 4C of this study – the multiple-staircase group, which showed significant transfer only at post-test. In the model, transfer either happens at mid-test or does not happen at all, assuming both stages of sequential training use the same staircase mode or are otherwise equally challenging. It may be the case that lack of transfer at mid-test seen in some double training studies may involve an additional inhibitory process (Zhang et al., 2013) that prevents transfer from increasing performance.

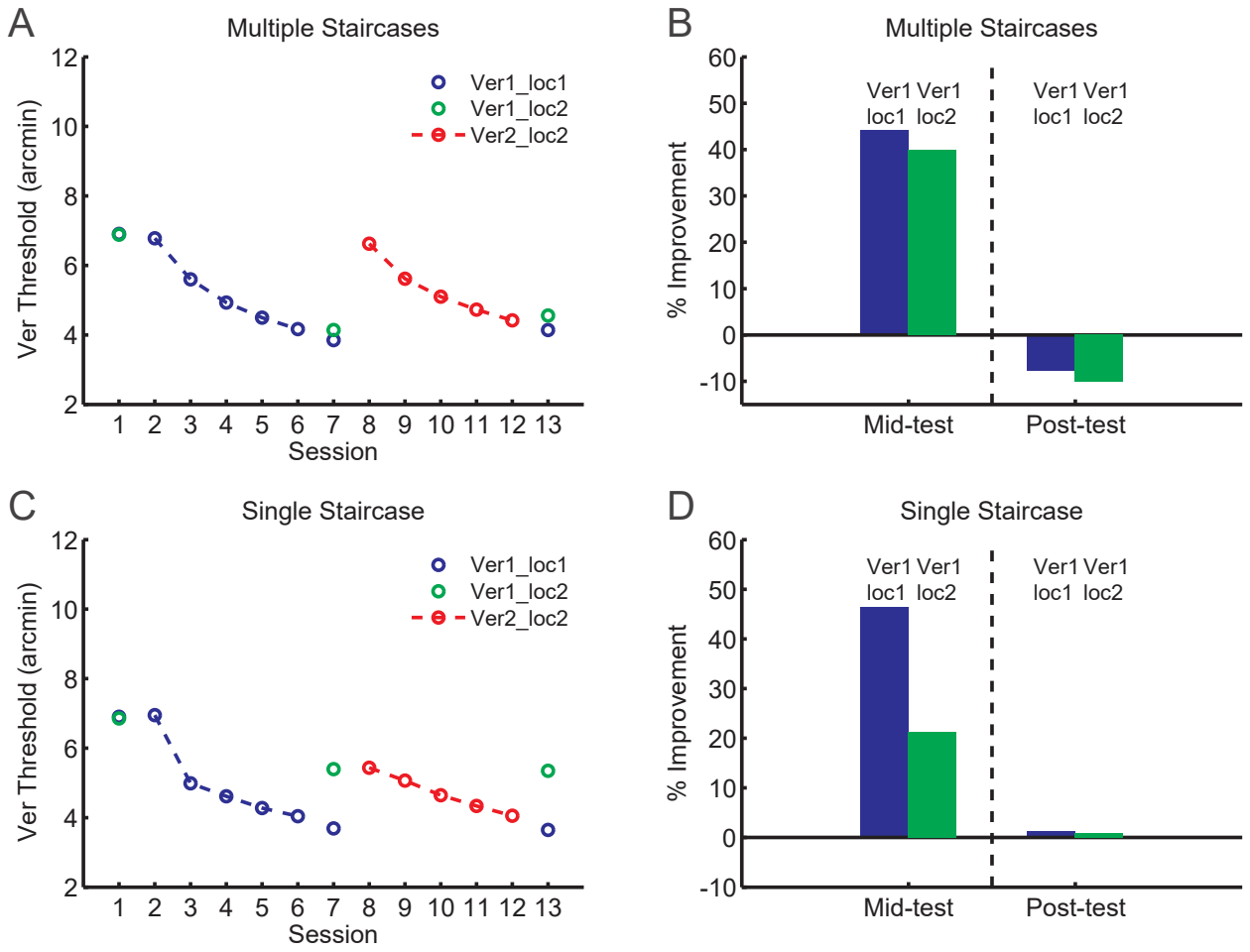

Figure 4: DPMM: results of simulations examining whether learning of a Vernier task of one orientation (Ver1_loc1) followed by learning of a Vernier task of an orthogonal orientation at another retinotopic location (Ver2_loc2) transfers to a Vernier task of the former orientation at the latter location (Ver1_loc2). Thresholds at the end of the 3 test (1, 7, 13) and 10 (2-6, 8-12) training sessions (A and C), as well as MPIs (B and D), are shown for the multiple and single staircase conditions, respectively. Error bars are omitted as they are smaller than the symbols themselves. Cf. Figure 4 of Hung & Seitz (2014) (reproduced in Figure E.14 of AppendixE).

3.2. Predictions

The best-fitting parameter values to the sequential training data of Hung & Seitz (2014) are used for all simulations that follow. Therefore the DPMM predicts the datasets described hereafter rather than fitting them.

3.2.1. Single vs multiple staircases in double training

Predictions for the single vs multiple staircase condition in the Ver1/Ver2 double training experiment of Hung & Seitz (2014) are shown in Figure 5 (cf. Figures 1 and 3 of their paper; these figures are also presented in Figure E.15 of AppendixE). The qualitative pattern of the predictions for the Vernier matches that of the data, although there are some quantitative differences. One is that in the data there is little learning of the orientation task, whereas in the model learning is just

as effective for this task as it is for Vernier. Another difference is that the Ver_loc2 threshold at post-test is higher than the respective threshold at mid-test of sequential training (MPI≈ 32% and 44% , respectively; Figure 4), i.e. there is less learning. This is due to the fact that the Vernier and orientation tasks affect different weights: the Vernier (horizontal or vertical) task mainly cause changes in the weights of the units selective for $-90°$ or $0°$, respectively, whereas the orientation task mainly causes changes in the weights of units selective for$-60°$ and $-30°$. However, the pattern that distinguishes the multiple and single staircase conditions is still present: there is nearly complete transfer in the former ($TI = 0.83$, Figure 5B) but a little more than half the transfer in the latter ($TI = 0.55$, Figure 5D).

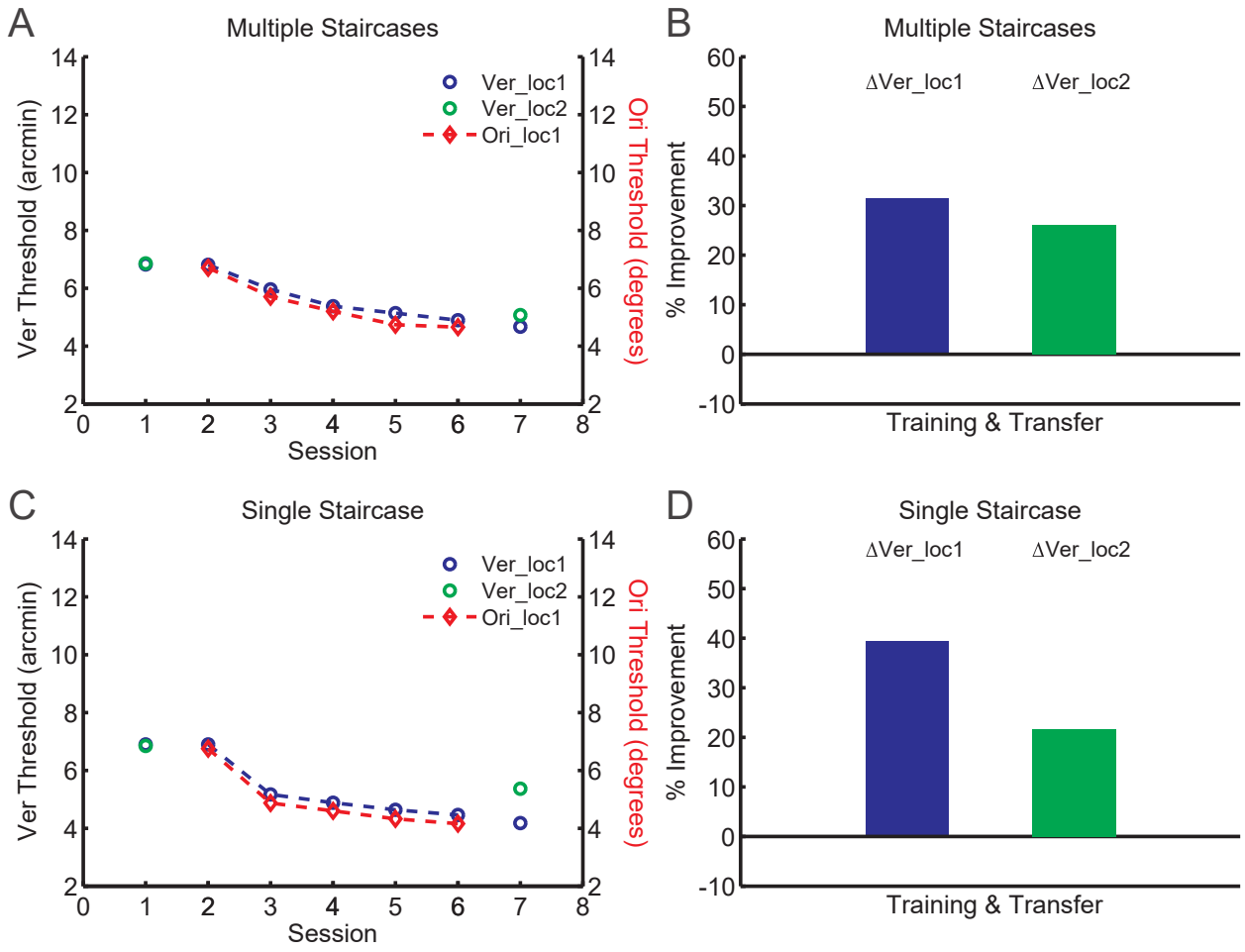

Figure 5: DPMM: results of simulations examining how learning of a concurrent Vernier and an orientation task of one orientation (Ver_loc1, Ori_loc1) transfers to a Vernier task of the same orientation at another retinotopic location (Ver_loc2). Thresholds at the end of the 2 test (1, 7) and 5 (2-6) training sessions (A and C), as well as MPIs (B and D), are shown for the multiple and single staircase conditions, respectively. Error bars are omitted as they are smaller than the symbols themselves. Cf. Figures 1 and 3 of Hung & Seitz (2014) (reproduced in Figure E.15 of AppendixE).

All parameter values used in the aforementioned simulations were also used in a control, single-

training simulation of the experiment of Hung & Seitz (2014) that examines location transfer of orientation discrimination. Figure 6 is analogous to the data shown in Figure 6 of their study. The DPMM in this case makes good quantitative predictions. In particular, it shows complete transfer of learning in the multiple staircase condition and partial transfer in the single staircase condition: MPI at loc2 (an indication of transfer) is around half the MPI at loc1. Also, the MPI at loc1 in the multiple staircase condition is slightly lower than the single-staircase MPI. These two relations are also seen in the data.

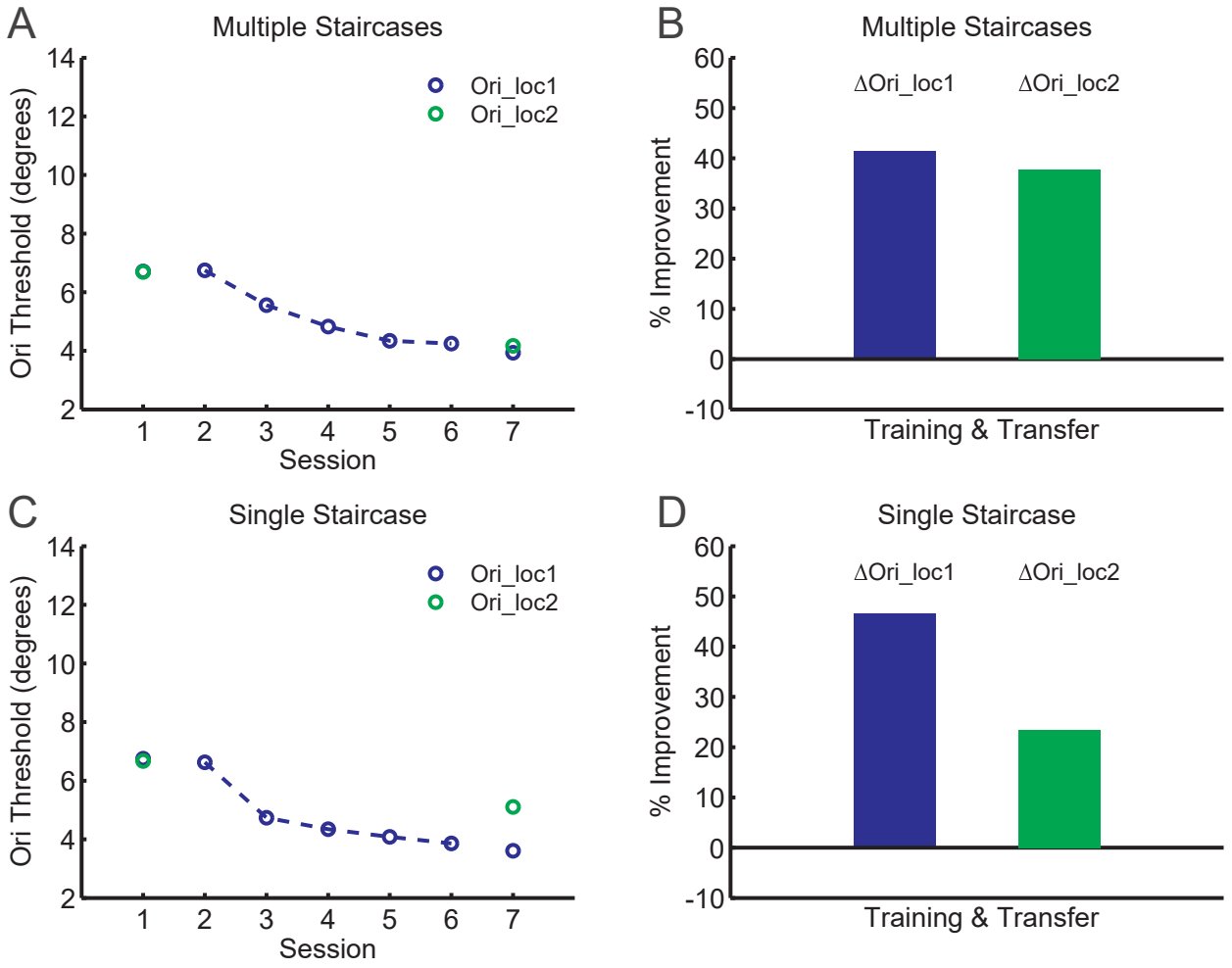

Figure 6: DPMM: results of simulations examining whether learning of an orientation task at one retinotopic location (Ori_loc1) transfers to another location (Ori_loc2). Thresholds at the end of the 2 test (1, 7) and 5 (2-6) training sessions (A and C), as well as MPIs (B and D), are shown for the multiple and single staircase conditions, respectively. Error bars are omitted as they are smaller than the symbols themselves. Cf. Figure 6 of Hung & Seitz (2014) (reproduced in Figure E.16 of AppendixE).

3.2.2. Single vs multiple staircases in other sequential training data

The DPMM makes a general prediction: sequential training with multiple short staircases cannot induce further transfer, beyond what single training can (i.e. beyond the mid-test), except in two

cases. One is the case of the short staircase terminating at relatively high threshold, which means fewer of the "easy" trials were actually easy. The other case is where the second training task is made easier than the first (i.e. easier than short 3/1 staircases). Interestingly, there is published data (Wang et al., 2012) that examines exactly this possibility, and the model can indeed account for it. The data in question is shown in Figures 2 and 3 of Wang et al. (2012), where the second training involves easy, suprathreshold tasks. After training on such easy tasks, the V4 representation ends up with a very high weight and location transfer occurs. All results in these two figures are predicted by the model. Only the data shown in Figure 4 of Wang et al. (2012) cannot be fully predicted: in the model, Con_ori2_loc2 learning would transfer to Ver_ori2_loc2, whereas in the data there was no such transfer.

3.2.3. Double training with Vernier/contrast

In the DPMM, the optimal weight vector for a Vernier task of a particular orientation is almost the same as the optimal vector for the contrast discrimination task of the same orientation. Therefore training on Vernier and contrast discrimination on one retinotopic location should result in transfer of Vernier to the other location. This is in conflict with the data of Wang et al. (2014), who found no transfer across retinotopic locations when the second trained task is contrast discrimination. First we wished to examine how well the model, with its current parameters, would fit their data (Figure 2c of their study). Thus the simulation was performed by keeping all previously mentioned model and experimental parameters the same. Remarkably, by choosing the width of the integration kernel for the contrast task so that the output activation across the range of contrast differences tested experimentally is similar in magnitude to the activation across the range of tested Vernier offsets, contrast threshold automatically matched those of human subjects, without the need for any adjustments to model parameters.

As we suspected, in the multiple-staircase condition the model cannot account for the absence of transfer of Vernier learning to an untrained location when double-training with contrast discrimination on a single location; transfer of Vernier learning was complete (Figure 7B). In fact, as far as the model is concerned, a Vernier task and contrast discrimination task of the same orientation are largely interchangeable. Therefore double training with Vernier and contrast for 8 blocks per session each in an alternating fashion results in almost as much learning and transfer of Vernier as single training on Vernier for 16 blocks per session. In other words, the Vernier learning curve in sessions 1 to 7 of Figure 7 is quantitatively very similar to the learning curve in sessions 1 to 7 of

the Vernier sequential training simulation (Figure 4).

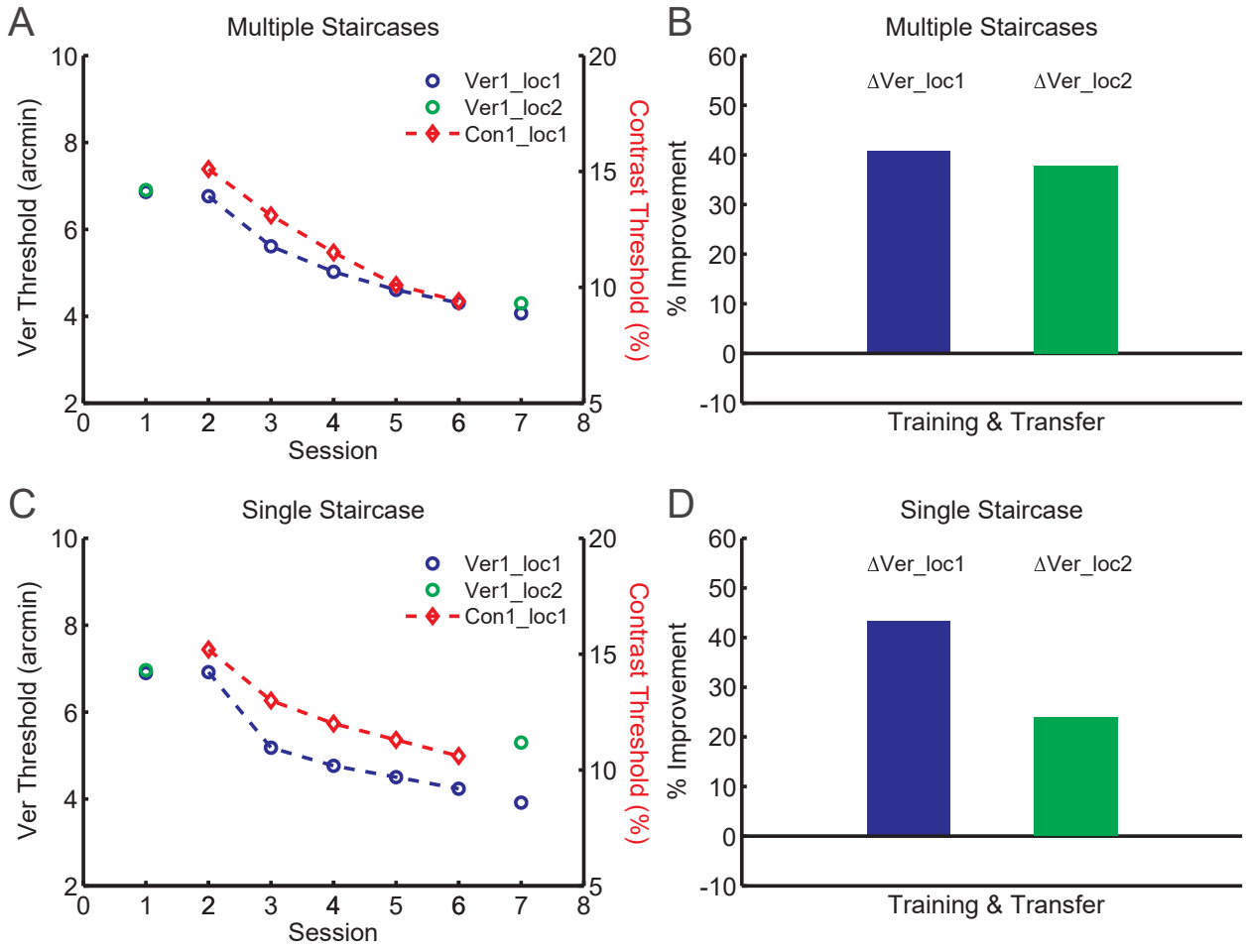

Figure 7: DPMM: results of simulations examining how learning of a concurrent Vernier and a contrast task of one orientation at the same retinotopic location (Ver1_loc1, Con1_loc1) transfers to a Vernier task of the same orientation at another retinotopic location (Ver1_loc2). Thresholds at the end of the 2 test (1, 7) and 5 (2-6) training sessions (A and C), as well as MPIs (B and D), are shown for the multiple and single staircase conditions, respectively. Error bars are omitted as they are smaller than the symbols themselves. Cf. Figure 2c of Wang et al. (2014) (reproduced in Figure E.17 of AppendixE).

However, although it is not clarified, Figure 2c in the study of Wang et al. (2014) shows a contrast stimulus with an orientation orthogonal to the Vernier stimulus. In this case, the model makes quantitatively approximately the same prediction as the one that follows (Ver1/Ver2 concurrent double training), which, as Figure 8 shows, still does not match the findings of Wang et al. (2014). Further data on this interesting differential effect of contrast reported by Wang et al. (2014) could help understand better the conditions under which it occurs.

3.2.4. Prediction: double training with Ver1/Ver2 in the same location

In light of the fact that the amount of learning of both Ver1 (vertical Vernier) and Ver2 (horizontal Vernier) in sequential training (Figure 4) is slightly smaller at post-test (MPI= 40%) than it is at mid-test, our hypothesis was that turning sequential training into double training while keeping all other parameters the same would yield similar results. Training with Ver1 increases the weights of units with preferred orientation close to vertical whereas Ver2 training increases the weights of units with preferred orientation close to horizontal. Therefore the MPI for Ver1_loc1 (training) and Ver1_loc2 (transfer) in double training should be the same as, if not less than, the respective MPIs in single/sequential training since the units that are informative for Ver1_loc1 are just a source of noise for Ver2_loc2 and vice-versa: V4 will end up with higher weights in both horizontal and vertically oriented units and therefore with higher noise for both tasks.

Figure 8 shows that this was indeed the case: as predicted, double training with two Vernier orientations in the same location resulted in less learning of both, although learning transferred to the second location as usual, i.e. there was near-complete transfer with multiple staircases and less transfer with a single staircase. There was no additional facilitation of transfer by training both Ver1 and Ver2 at the same location – no "piggybacking" effect such as the one found by Wang et al. (2014).

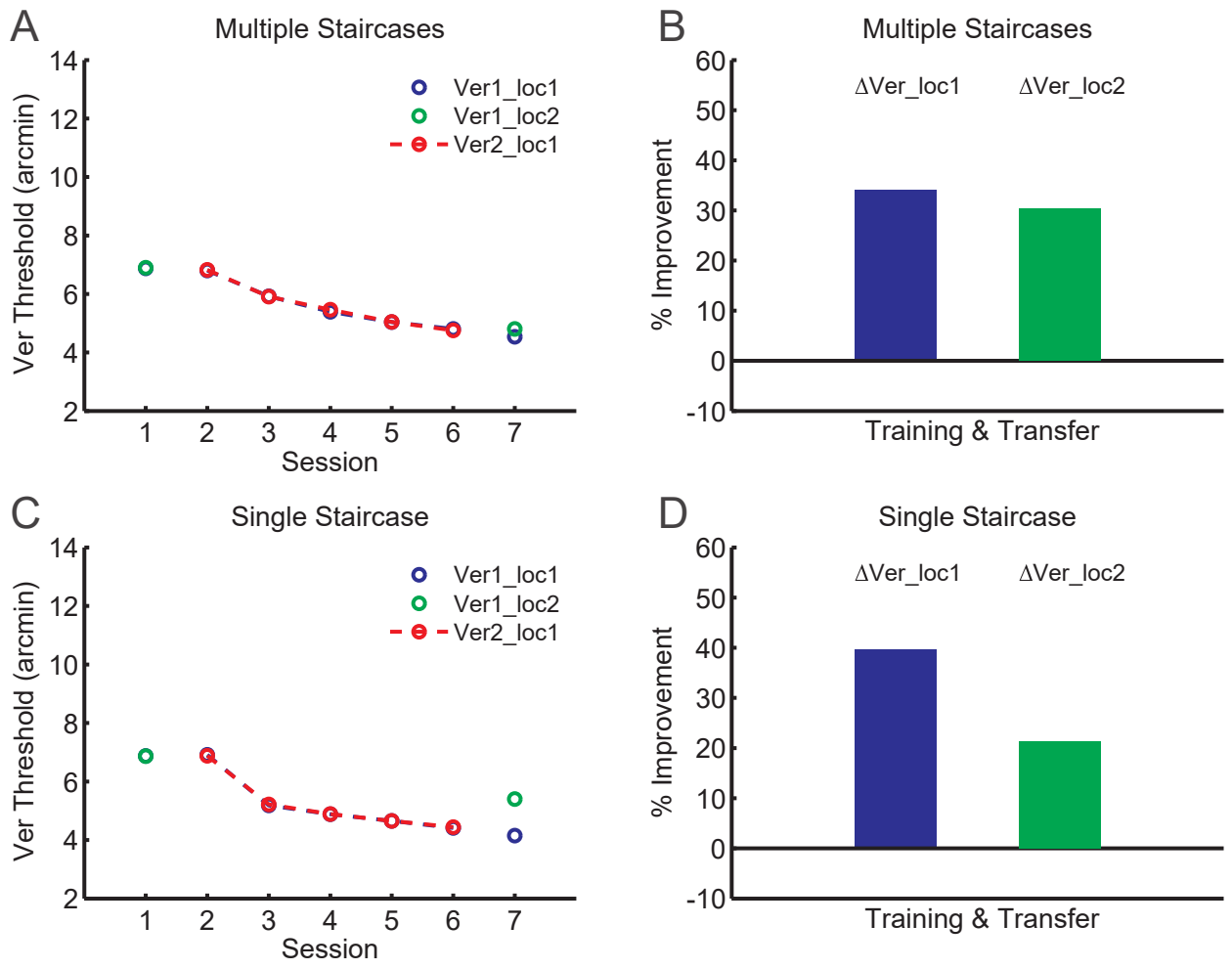

Figure 8: DPMM: results of simulations examining how concurrent learning of two Verniers of perpendicular orientations at the same retinotopic location (Ver1_loc1, Ver2_loc1) transfers to one of them at another retinotopic location (Ver1_loc2). Thresholds at the end of the 2 test (1, 7) and 5 (2-6) training sessions (A and C), as well as MPIs (B and D), are shown for the multiple and single staircase conditions, respectively. Error bars are omitted as they are smaller than the symbols themselves. The learning curves of Ver1_loc1 and Ver2_loc1 are similar enough so that the former overlaps the latter almost perfectly and renders it invisible.

3.2.5. Prediction: Vernier training with 2 down, 1 up staircase

It has become apparent that, within the DPMM, training with multiple short staircases results in nearly complete retinotopic location transfer because short staircases contain several easy, suprathreshold trials and thus the performance of V4 is high enough so that its contribution to the perceptual decision is high. But is it the number of easy trials in the staircase that makes it different to a single staircase or is it the proportion of easy trials? Hung & Seitz (2014) hypothesized that it is the latter rather than the former: if a staircase contains a high enough proportion of near-threshold trials, the mere presence of any particular number of easy trials should not induce significant transfer. The DPMM, which computes the contribution on V4 based on its performance over the last

several trials (it has a rate parameter 0.0024, yielding a time constant of 416 exponentially discounted trials), supports in theory the hypothesis of Hung & Seitz (2014). Here we decided to test this theory by modifying the staircase rule from the usual 3-down, 1-up that converges to 79.4% threshold and is used in all experiments (Wang et al., 2012; Hung & Seitz, 2014; Wang et al., 2014), to the 2-down, 1-up rule, which converges to 70.7% performance, i.e. at a lower threshold. This way, the number of easy trials (that correspond to at least 79.4% performance) remains exactly the same but the proportion of harder trials increases (since there is now a number of trials corresponding to performances between 70.7% and 79.4%. According to the hypothesis of Hung & Seitz (2014), there should be significantly less location transfer, even when training with multiple short staircases.

Figure 9 shows this was indeed the case. The 2/1 rule caused the difference between multiple staircases and single staircase to almost vanish, with the former resulting in $TI = 0.52$ and the latter in $TI = 0.44$, i.e. both staircases now resemble the single staircase of previous simulations.

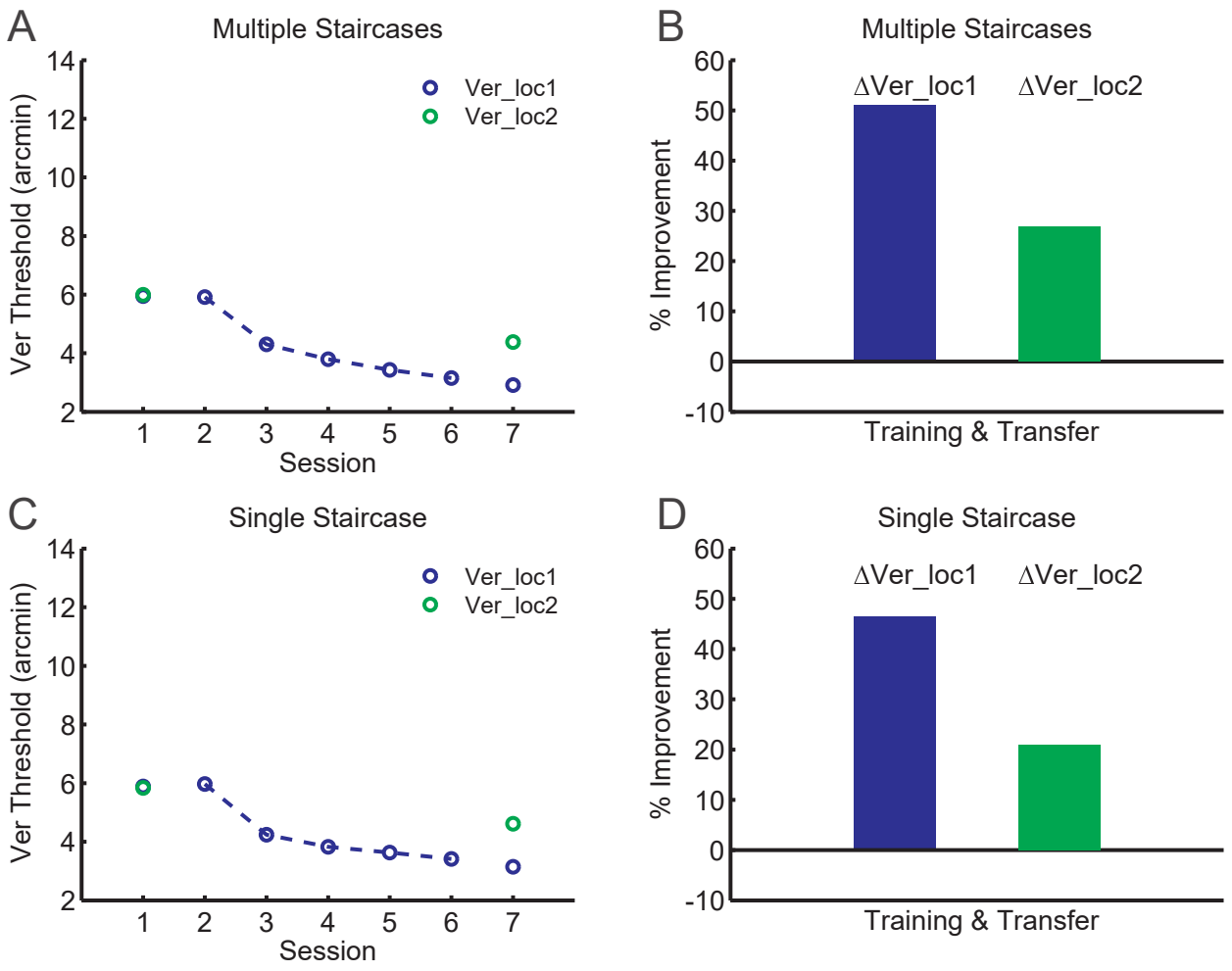

Figure 9: DPMM: results of simulations examining whether learning of a Vernier at one retinotopic location (Ver_loc1) transfers to another location (Ver_loc2). Thresholds at the end of the 2 test (1, 7) and 5 (2-6) training sessions (A and C), as well as MPIs (B and D), are shown for the multiple and single staircase conditions, respectively. Note that the staircases used in these simulations obeyed the 2-down, 1-up rule, which converges to 70.7% performance. Error bars are omitted as they are smaller than the symbols themselves.

4. Discussion

The reweighting model of Dosher & Lu (1998), in the last almost 20 years, has evolved into the AHRM/IRT and by now has been able to provide quantitative accounts for a multitude of datasets in perceptual learning. The reweighting hypothesis, which posits that much of perceptual learning can be explained by processes whereby high-level areas adjust their readouts from early sensory areas, has recently been invoked to explain the transfer of learning across retinotopic locations (Dosher et al., 2013). The simple architecture of the model, which is essentially feed-forward with a decision unit that linearly combines sensory inputs, has demonstrated how far a "bare essentials" model can go towards explaining diverse findings in perceptual learning and is further testament

to the power of the reweighting hypothesis. Here, with the DPMM, we extended the IRT model beyond its one-layer feed-forward architecture to explore the conditions under which transfer across retinotopic locations might occur, taking into account recent experimental findings (Hung & Seitz, 2014; Wang et al., 2014).

We found that multiple short staircases, which contain a larger proportion of easy trials, increase the signal from V4 and thus its weight in the perceptual decision. At the same time, these easy trials result in activations of greater magnitude (whether positive or negative) compared to the activations of hard trials (where the network output is near zero) and thus the Hebbian learning rule increases the weights of the informative units more than it does in single-staircase mode. This largely compensates for the naturally higher selectivity of V1 (which is less noisy and more precise than V4) so that the final thresholds under both staircase modes are very similar, as seen in the data of Hung & Seitz (2014).

4.1. Related work

Overall, the DPMM provides a good quantitative account of the important aspects of the data under consideration. Its predecessor (Talluri et al., 2015) is the only one, to date, that addresses the experimental findings of Hung & Seitz (2014). However, the model is rather restricted in its applicability and cannot address other double-training findings (such as those of Wang et al., 2014) and possibly earlier PL findings that the AHRM/IRT can account for. Furthermore, the model of Talluri et al. (2015) uses the delta learning rule, which is neither purely associative nor demonstrably biologically plausible. In the DPMM weights are learned in an associative manner, augmented by feedback (either in the way feedback augments associative learning in AHRM or in the way the contribution of V4 is proportional to its recent performance). Lastly, the model of Talluri et al. (2015) includes an *ad hoc* criterion for learning the weights of connections from each representation layer (V1 and V4) to the decision unit, by computing "confidence" – a quantity proportional to the magnitude of the network output. The DPMM, on the other hand, is designed on normative principles, i.e. it directly tracks performance in V4 in an incremental, biologically plausible way that requires no memory, and it uses that performance estimate to adjust the contribution of V4 to the decision based on a performance/transfer trade-off controlled by two parameters.

As mentioned in the beginning, the DPMM is conceptually very similar to mixture-of-experts (MoE) models (Jacobs et al., 1991b; Jordan & Jacobs, 1991; Hampshire & Waibel, 1992; Jacobs, 1997). In both DPMM and MoE models, the contribution of the individual experts (here V1 and V4)

are gated according to some criterion. In DPMM this criterion is a trade-off between retinotopic location invariance of learning and performance; in MoE models the criterion is purely performance. Importantly, however, the gating in both models is *multiplicative*, as explained in 2.2.6. In recent years there has been renewed interest in multiplicative gating (MG) in the context of recurrent neural networks (RNN) used in sequence learning tasks, such as character language modeling, speech recognition and text generation (Chung et al., 2015; Krause et al., 2016; Zhang et al., 2016; Wu et al., 2016). Wu et al. (2016) show that their form of MG (Multiplicative Integration, MI) can seen as a nonlinear, and more flexible, extension of Hidden Markov Models and that several popular RNNs, such as the long short-term memory (LTSM) of Hochreiter & Schmidhuber (1997), can be optimized more easily and show better generalization of learning when equipped with MI. Incidentally, our model exhibits both of these features, although generalization in the present work is implicitly "unlocked" rather than emerging (given that V4 is already assumed to be retinotopic location-invariant).

Despite their similarities, the DPMM differs from MoE models in a number of ways. First, biological plausibility is not a concern in MoEs; for example, MoEs are trained with the delta rule. Second, in canonical MoE models, gating is based on the expert network's outputs, i.e. after the activation function has been applied to the input-weight dot product; whereas in the DPMM this occurs prior to the activation function (Eq. 14). In this sense, the DPMM gating resembles the model of Wu et al. (2016) mentioned earlier, in which the multiplicative operation is also performed prior to the activation function. Third, whereas in the DPMM the V1/V4 weighting mechanism produces a (deterministic) mixture of the individual network responses, in canonical MoEs the gating weight is actually the probability of selecting a particular expert (i.e. gating involves a stochastic switch that gives all weight to one expert at a time). In this sense, our model is more similar to that of Hampshire & Waibel (1992) than to the canonical MoEs (Jacobs et al., 1991b; Jacobs, 1997). Lastly, the coupling between the weight of V4 and its performance is direct (and its strength is controlled by the slope of the sigmoid in Eq. 12), whereas in the canonical MoEs, the coupling is indirect and generally weaker. We note, however, that MoE models in literature are trained in a variety of ways, resulting in different gating behaviours. Notably, there is a instantiation of a MoE model in which the error function consists of an exponentially weighted average of past errors that is compared to the current error (Jacobs et al., 1991a). Our Eq. 11 bears striking similarities to Eqs. 4 and 5 in that work, where multiplicative gating is realized more

directly.

4.2. Limitations

Although the DPMM provides a reasonably accurate quantitative account of most of the data of Hung & Seitz (2014), there is data from other labs that it cannot fit. Two interesting cases appear in the same study (Wang et al., 2014). One is the differential effect of contrast, already discussed in 3.2.3. This effect, if robust, will require an additional feature in the model, such as multiple representations, multiple weight vectors or representation modification. Another is the differential effect of ordering of Vernier and Orientation tasks in a sequential training experiment, where location transfer of Vernier learning at post-test occurs if Vernier training is followed by orientation discrimination training but not vice versa (Wang et al., 2014). We suspect that for ordering effects to be relevant, the weight update rule will have to depart from the Hebbian equation and include some form of weight decay or an error-driven update rule (such as the delta rule). Oja's learning rule (Oja, 1982) is a potentially suitable candidate and there are theoretical reasons to believe it is biologically realistic. However, Oja's rule or any rule that causes weight decay at time scales comparable to the duration of a perceptual learning experiment would also render the model incapable of learning more than one task.

4.3. Predictions and future directions

A central assumptions in the model, without which there would be no transfer or interference between tasks, is that the same connections and weights are used for all three tasks. This is just one of a number of possibilities however; under the task analysis of Petrov et al. (2005), the DPMM belongs in one of four categories – the one in which both the representation pool and the task-dependent decision structure are shared. Another category, for example, is that different sets of connections to the same representation pool (i.e. separate decision units) exist for each task and/or stimulus. However, the data on disruption (Seitz et al., 2005; Zhang et al., 2008; Yotsumoto et al., 2009) and transfer (Webb et al., 2007; Xiao et al., 2008; Jeter et al., 2009; Zhang et al., 2010; Wang et al., 2012; Dosher et al., 2013; Wang et al., 2014) of learning imply at least some form of overlap of decision structures or, alternatively, modification of shared representations. We emphasize that the DPMM, with its minimal assumptions regarding the multiplicity of representations and decision structures, is more of a proof of concept exploring what can be achieved with a single representation

subsystem[2] equipped with MG, given that the three perceptual tasks we simulated used identical stimuli. Future directions would be the inclusion of multiple representation subsystems and the same or a modified gating mechanism; such a model, which would be closer to a MoE in terms of input/task partitioning, would likely be able to handle a wider variety of tasks and stimuli.

It should be stressed that in both the DPMM and its predecessor, the IRT, learning consists of changes in the weights of connections between representation and decision areas for both location-dependent (V1) and location-independent (V4) representations. In other words, the differences caused by the different staircase modes are not differences in learning – they are differences in contribution to decision areas. Therefore the model predicts that even after prolonged training at threshold (which minimizes the contribution of V4), just a *few* blocks that contain a large proportion of easy trials (e.g. a couple of short staircases) are sufficient to induce transfer of learning across retinotopic locations. In other words, *learning occurs everywhere*, it just needs to be unlocked with the appropriate experimental manipulation. This is a prediction that can be easily verified experimentally and would provide a good test of the model. Furthermore, the model predicts the location specificity so often seen in classical studies of PL. These studies have typically used demanding tasks, with staircases that converge to less than 80% accuracy (even if multiple short ones are used during training, as we show in 3.2.5), or other methods, such as that of constant stimuli, that contain a relatively small proportion of easy trials (or even only subthreshold trials, as in Herzog & Fahle, 1997). In such cases, the model assigns a very low weight on V4.

The present extension to the AHRM/IRT, and especially the novel dynamic V1/V4 weighting, extends its applicability to a broader set of findings in perceptual learning. However, it is important to note that the attribution of learning in the model to areas V1 and/or V4 is speculative and is just one of a number of instantiations of where learning could be attributed in the brain. Likewise, while the stated framework is as a read-out model, this does not necessarily mean that learning actually occurs in the weights to decision structures and it could be taking place in part in the representations, or decision areas. Moreover, although dynamic V1/V4 weighting system has been intentionally specified in abstract terms without committing to a particular architecture, the multiplicative gating introduced in the model almost necessarily implies feedback connections to V1 and V4. One (but by no means the only) biologically plausible instantiation would consist

[2]Part of this pool is location-dependent, e.g. V1 units, and part of it is independent, e.g. V4 units; the point is that the same representations process all three tasks/stimuli.

of a unit for the running average (which would work similarly to the bias unit in the AHRM/IRT and the DPMM, although for a different purpose) that receives input from a V4 "decision unit" and the feedback unit and projects "back" to an excitatory unit and to two (at least) inhibitory units. These last three units could be near/at the representation level: the excitatory one would connect to every V4 unit and each inhibitory one to every V1 unit of the respective retinotopic location. The weights of the connections to and from the three units could be predetermined (perhaps during familiarization with the task) but remain fixed throughout a simulation. Such an architecture would effectively implement multiplicative gating of the V1 and V4 units. An interesting elaboration of the DPMM could implement these units explicitly; this is left as future work.

Regardless of the exact neural instantiation, this model is an important step towards understanding under what conditions perceptual learning can generalize (across retinotopic locations and/or tasks). Further understanding of how the visual system is able to cope with the diversity of visual stimuli in the world can help design more efficient protocols for visual rehabilitation (Polat, 2009; Deveau et al., 2013) or enhancement in healthy populations (Deveau et al., 2014).

5. Acknowledgements

We thank the reviewers for their insightful and useful comments during revision of the manuscript.

Funding: this work was supported by the National Institutes of Health [grant number 1R01EY023582].

References

Ahissar, M., & Hochstein, S. (1993). Attentional Control of Early Perceptual Learning. *Proceedings of the National Academy of Sciences*, *90*, 5718–5722.

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, *8*, 457–464.

Bienenstock, E., Cooper, L., & Munro, P. (1982). Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience*, *2*, 32–48.

Chung, J., Gülçehre, C., Cho, K., & Bengio, Y. (2015). Gated feedback recurrent neural networks. In *ICML* (pp. 2067–2075).

Crist, R., Kapadia, M., Westheimer, G., & Gilbert, C. (1997). Perceptual learning of spatial localization: specificity for orientation, position, and context. *Journal of Neurophysiology*, *78*, 2889–2894.

Deveau, J., Lovcik, G., & Seitz, A. R. (2013). The therapeutic benefits of perceptual learning. *Current trends in neurology*, *7*, 39.

Deveau, J., Ozer, D. J., & Seitz, A. R. (2014). Improved vision and on-field performance in baseball through perceptual learning. *Current Biology*, *24*, R146–R147.

Dosher, B., & Lu, Z. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proceedings of the National Academy of Sciences*, *95*, 13988–13993.

Dosher, B. A., Jeter, P., Liu, J., & Lu, Z.-L. (2013). An integrated reweighting theory of perceptual learning. *Proceedings of the National Academy of Sciences*, *110*, 13678–13683.

Fahle, M. (1997). Specificity of learning curvature, orientation, and vernier discriminations. *Vision Research*, *37*, 1885–1895.

Fahle, M., & Morgan, M. (1996). No transfer of perceptual learning between similar stimuli in the same retinal position. *Current Biology*, *6*, 292–297.

Hampshire, J. B., & Waibel, A. (1992). The meta-pi network: Building distributed knowledge representations for robust multisource pattern recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *14*, 751–769.

Herzog, M., & Fahle, M. (1997). The role of feedback in learning a vernier discrimination task. *Vision Research*, *37*, 2133–2141.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*, 1735–1780.

Huang, C.-B., Lu, Z.-L., & Dosher, B. A. (2012). Co-learning analysis of two perceptual learning tasks with identical input stimuli supports the reweighting hypothesis. *Vision research*, *61*, 25–32.

Hung, S.-C., & Seitz, A. R. (2014). Prolonged training at threshold promotes robust retinotopic specificity in perceptual learning. *The Journal of Neuroscience*, *34*, 8423–8431.

Jacobs, R. A. (1997). Bias/variance analyses of mixtures-of-experts architectures. *Neural computation*, *9*, 369–383.

Jacobs, R. A., & Jordan, M. I. (1993). Learning piecewise control strategies in a modular neural network architecture. *IEEE Transactions on Systems, Man, and Cybernetics*, *23*, 337–345.

Jacobs, R. A., Jordan, M. I., & Barto, A. G. (1991a). Task decomposition through competition in a modular connectionist architecture: The what and where vision tasks. *Cognitive Science*, *15*, 219–250.

Jacobs, R. A., Jordan, M. I., Nowlan, S. J., & Hinton, G. E. (1991b). Adaptive mixtures of local experts. *Neural computation*, *3*, 79–87.

Jeter, P. E., Dosher, B. A., Petrov, A., & Lu, Z.-L. (2009). Task precision at transfer determines specificity of perceptual learning. *Journal of Vision*, *9*, 1–1.

Jordan, M. I., & Jacobs, R. A. (1991). Hierarchies of adaptive experts. In *NIPS* (pp. 985–992).

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences*, *88*, 4966–4970.

Krause, B., Lu, L., Murray, I., & Renals, S. (2016). Multiplicative lstm for sequence modelling. *arXiv preprint arXiv:1609.07959*, .

Law, C., & Gold, J. (2008). Neural correlates of perceptual learning in a sensory-motor but not a sensory cortical area. *Nature Neuroscience*, *11*, 505–513.

Lu, Z.-L., Liu, J., & Dosher, B. A. (2010). Modeling mechanisms of perceptual learning with augmented hebbian re-weighting. *Vision research*, *50*, 375–390.

Oja, E. (1982). Simplified neuron model as a principal component analyzer. *Journal of mathematical biology*, *15*, 267–273.

Petrov, A., Dosher, B., & Lu, Z. (2005). The dynamics of perceptual learning: An incremental reweighting model. *Psychological Review*, *112*, 715–743.

Petrov, A., Dosher, B., & Lu, Z. (2006). Perceptual learning without feedback in non-stationary contexts: Data and model. *Vision Research*, *46*, 3177–3197.

Poggio, T., Fahle, M., & Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. *Science, New Series*, *256*, 1018–1021.

Polat, U. (2009). Making perceptual learning practical to improve visual functions. *Vision research*, *49*, 2566–2573.

Schoups, A., Vogels, R., & Orban, G. (1995). Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. *Journal of Physiology*, *483*, 797–810.

Schoups, A., Vogels, R., Qian, N., & Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature*, *412*, 549–552.

Seitz, A., Yamagishi, N., Werner, B., Goda, N., Kawato, M., & Watanabe, T. (2005). Task-specific disruption of perceptual learning. *Proceedings of the National Academy of Sciences*, *102*, 14895–14900.

Shiu, L., & Pashler, H. (1992). Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Perception and Psychophysics*, *52*, 582–588.

Sotiropoulos, G., Seitz, A. R., & Seriès, P. (2011). Perceptual learning in visual hyperacuity: A reweighting model. *Vision research*, *51*, 585–599.

Sowden, P., Rose, D., & Davies, I. (2002). Perceptual learning of luminance contrast detection: specific for spatial frequency and retinal location but not orientation. *Vision Research*, *42*, 1249–1258.

Talluri, B. C., Hung, S.-C., Seitz, A. R., & Seriès, P. (2015). Confidence-based integrated reweighting model of task-difficulty explains location-based specificity in perceptual learning. *Journal of vision*, *15*, 17–17.

Wang, R., Zhang, J.-Y., Klein, S. A., Levi, D. M., & Yu, C. (2012). Task relevancy and demand modulate double-training enabled transfer of perceptual learning. *Vision research*, *61*, 33–38.

Wang, R., Zhang, J.-Y., Klein, S. A., Levi, D. M., & Yu, C. (2014). Vernier perceptual learning transfers to completely untrained retinal locations after double training: A "piggybacking" effect. *Journal of vision*, *14*.

Webb, B., Roach, N., & McGraw, P. (2007). Perceptual learning in the absence of task or stimulus specificity. *PLoS ONE*, *2*, e1323. doi:10.1371/journal.pone.0001323.

Weiss, Y., Fahle, M., & Edelman, S. (1993). Models of perceptual learning in vernier hyperacuity. *Neural Computation*, *5*, 695–718.

Wu, Y., Zhang, S., Zhang, Y., Bengio, Y., & Salakhutdinov, R. R. (2016). On multiplicative integration with recurrent neural networks. In *Advances in Neural Information Processing Systems* (pp. 2856–2864).

Xiao, L., Zhang, J., Wang, R., Klein, S., Levi, D., & Yu, C. (2008). Complete transfer of perceptual learning across retinal locations enabled by double training. *Current Biology*, *18*, 1922–1926.

Yang, T., & Maunsell, J. (2004). The effect of perceptual learning on neuronal responses in monkey visual area V4. *Journal of Neuroscience*, *24*, 1617–1626.

Yotsumoto, Y., Chang, L., Watanabe, T., & Sasaki, Y. (2009). Interference and feature specificity in visual perceptual learning. *Vision Research*, *49*, 2611–2623.

Zhang, G.-L., Cong, L.-J., Song, Y., & Yu, C. (2013). Erp p1-n1 changes associated with vernier perceptual learning and its location specificity and transfer. *Journal of vision*, *13*, 19–19.

Zhang, J., Kuai, S., Xiao, L., Klein, S., Levi, D., & Yu, C. (2008). Stimulus coding rules for perceptual learning. *PLoS Biology*, *6*, e197.

Zhang, J.-Y., Zhang, G.-L., Xiao, L.-Q., Klein, S. A., Levi, D. M., & Yu, C. (2010). Rule-based learning explains visual perceptual learning and its specificity and transfer. *The Journal of Neuroscience*, *30*, 12323–12328.

Zhang, S., Wu, Y., Che, T., Lin, Z., Memisevic, R., Salakhutdinov, R. R., & Bengio, Y. (2016). Architectural complexity measures of recurrent neural networks. In *Advances in Neural Information Processing Systems* (pp. 1822–1830).

Appendix

AppendixA. Learning rule

The reason that the weight decay that the original learning rule entails has not been a problem in previous AHRM literature is that the model has only been used to simulate single tasks – typically orientation discrimination using gratings[3], such as the stimuli in Petrov et al., 2005. In fact, the authors considered weight decay an advantage for their task and stimuli:

"*The weight-bounding mechanism approximates the desired effect indirectly by penalizing the highly variable units relative to the less variable units. For instance, assume a weight* $w > 0$ *is updated in opposite directions* u *and* $-u$ *on consecutive trials. Because of the saturating nonlinearity in Equation 13, the negative update has greater impact than does the positive update, resulting in an overall regression toward the neutrality point* $w = 0$*. The more variable the updates – for example,* $\pm 2u$ *instead of* $\pm u$ *– the stronger the regressive effect. An inverse relationship between connection strength and activation variability emerges, approximating the optimal solution.*"

In Appendix B of the same study, the authors state:

"*The weight-bounding Equation 13 imposes the additional constraint that the norm of the weight vector is approximately constant:* $\|\mathbf{w}\| \approx n$ *(*≈ 1 *in the simulations). Thus, the decision noise cannot be eliminated by indiscriminate strengthening of the bottom-up connections.*"

Firstly, simulations with an orientation discrimination task show that this statement is not generally true: the weights of the most informative channels increase in absolute value (positive weights become more positive and negative ones more negative) and therefore the 2-norm of the weight vector steadily increases. It is the sum of the elements of the weight vector that stays approximately constant (zero). More importantly however, there is no good reason that the 2-norm stay constant throughout training, as long as the weights stay well below their maximum values throughout the simulation of psychophysical experiments of the maximum plausible duration (number of trials). Furthermore, the new sliding average bias calculation method described earlier is an approximation of the BCM rule (Bienenstock et al., 1982). This rule has been shown to stabilize Hebbian learning and, as mentioned earlier, we observed this increased stability in our simulations to the point that it did not become an issue with the learning rates suitable for these simulations, even with large

[3] A notable exception is the AHRM implementation of Huang et al., 2012, who used an entirely different representation system to model 3-dot alignment and bisection tasks.

amounts of noise in the system.

The original weight update rule has the advantage that it gradually decreases the weights of units that are not responsive to a stimulus of a particular orientation and thus only contribute noise to the output. This weight decay, however, also prevents the model from learning two tasks concurrently, especially if the two tasks correspond to two different (orthogonal) orientations of the same stimulus. In other words, there is tradeoff between noise reduction in (and therefore optimization of) a particular task and the ability to learn a variety of different tasks.

AppendixB. Integration kernels and local energy maps

For each preferred orientation and spatial frequency there are two units. For the Vernier and contrast tasks, the second unit has a kernel that is symmetric to the first about the horizontal axis (in other words, the second kernel is a vertically "flipped" version of the first). For example, for the Vernier task, one kernel has the integration centers on the top right and bottom left and the other kernel on the top left and bottom right. For the orientation task, where the kernel is at the center, the second unit has a null kernel, i.e. it is inactive. The units with the flipped (or, in the case of the orientation task, null) kernels are assigned equal but opposite weights to those of the respective units with the original kernels. The reason for this is that this way the weighted input $\mathbf{w} \cdot \mathbf{X}$ (and thus the output of the decision unit) is a monotonic function of stimulus offset, has the same magnitude but opposite sign for equal but opposite offsets and has a value of zero at zero offset. To see this, let $A_{\theta,f}(o)$ be the average (i.e. excluding representation noise) response of the unit of a particular orientation (θ) and spatial frequency (f) preference to a stimulus of Vernier offset (or contrast difference, in the case of the contrast task) o (in its signed form). Let $B_{\theta,f}(o)$ the response of the respective unit with the flipped kernel. Because of both horizontal and vertical symmetry, $A_{\theta,f}(o) = B_{\theta,f}(-o)$. If the weights assigned to these two units are opposite, then the weighted sum of these units is

$$w_{\theta,f}A_{\theta,f}(o) - w_{\theta,f}B_{\theta,f}(o) = w_{\theta,f}(A_{\theta,f}(o) - A_{\theta,f}(-o))$$

It follows that for a zero-offset stimulus ($o = 0$) this sum is zero. In the case of the orientation task, which is a 2-IFC task, the network output is already zero when $o = 0$ and is symmetric about zero (equal but opposite offsets result in equal but opposite $\mathbf{w} \cdot \mathbf{X}$). In fact, the second set of units results in a symmetry between the Vernier and contrast stimuli on one hand and the orientation stimulus on the other hand. This is because the orientation stimulus has a reference orientation ($-45°$) located

in the middle of the orientation selectivity range (-90° to 0°) and thus units on either side of the stimulus are responsive to it, compared to the Vernier and contrast stimuli, which have vertical or horizontal gratings and thus only half the units are responsive to it (the ones with vertical or horizontal orientation selectivity). Furthermore, for the orientation stimulus, the units on either side of the reference orientation have the opposite monotonicity as a function of orientation difference $\Delta\theta$ (the units counterclockwise to the reference orientation are decreasing functions of $\Delta\theta$ whereas the units clockwise are increasing functions). The second set of units with the flipped kernels provide exactly this feature for the Vernier and contrast stimuli. Thus the exact same weights can be used for all three tasks.

The kernels for the 3 tasks were hardcoded in the model for parsimony, so that the same $5\times7\times2$ representation pool can be used for all tasks and stimuli. In a more biologically realistic implementation, there would be several different representation units with a variety of kernel configurations, and the network would learn to pay attention to those representations with the appropriate kernels for the task. This, however, would increase the complexity and the computational cost of the model without adding anything to the main points in the present work.

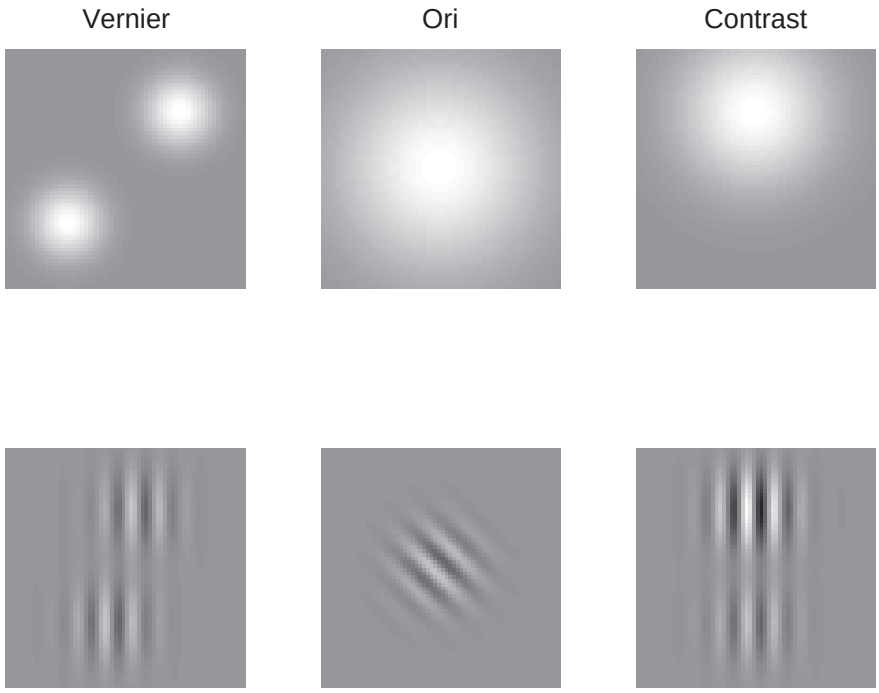

Figure B.10: Integration kernels (top row) and stimuli (bottom row) for the 3 tasks in the present simulations. Each kernel is applied to the input representation of each unit (see p742 in Petrov et al., 2005).

Apart from the location of the integration kernels, another parameter that is important is the

full-width-at-half-height h_r of the kernel (another way of expressing the standard deviation of the Gaussian component). In their simulations, Petrov et al. (2005) found that changing h_r from 2 to 1 d.v.a. did not significantly alter their results; we found that this parameter greatly affects the sensitivity of the representation units at values below 1 d.v.a. In this range, smaller values lead to higher precision of the representation subsystem but at the cost of asymmetries between stimuli with equal but opposite offsets. Figure B.10 shows integration kernels used in the present simulations, corresponding to $h_r = 2$ for the orientation task (which is the value used in all previous AHRM studies that involve orientation discrimination) but lower values for the Vernier and contrast discrimination tasks (Table 2). These values were chosen so that the 3 tasks result in similar magnitudes of output activation across the range of stimulus difficulties (offset, orientation or contrast difference), given the same initial weight vector. While the resulting kernels for the Vernier task have a slightly smaller diameter than that of the visible part of the gratings, such a value seems appropriate if one looks at how each representation unit "sees" the input image. Figure B.11 (also see Figure 6 in Petrov et al., 2005) shows local energy maps of the 5×7 representation units for the Vernier stimulus of Figure B.10, bottom left. The diameter of the largest blob in the most active energy maps is in fact comparable to the diameter of the respective integration kernel for that task. Note that for display purposes the figures of the maps are normalized by the minimum and maximum values of the particular map. For example, the white regions in the top left map correspond to much lower energies than the white regions in the bottom right map. Were it not for this normalization, most of these maps would look entirely black as the local energy is near zero. The pooled energy of each representation unit for the easiest (greatest offset) stimulus is shown in the title of each map.

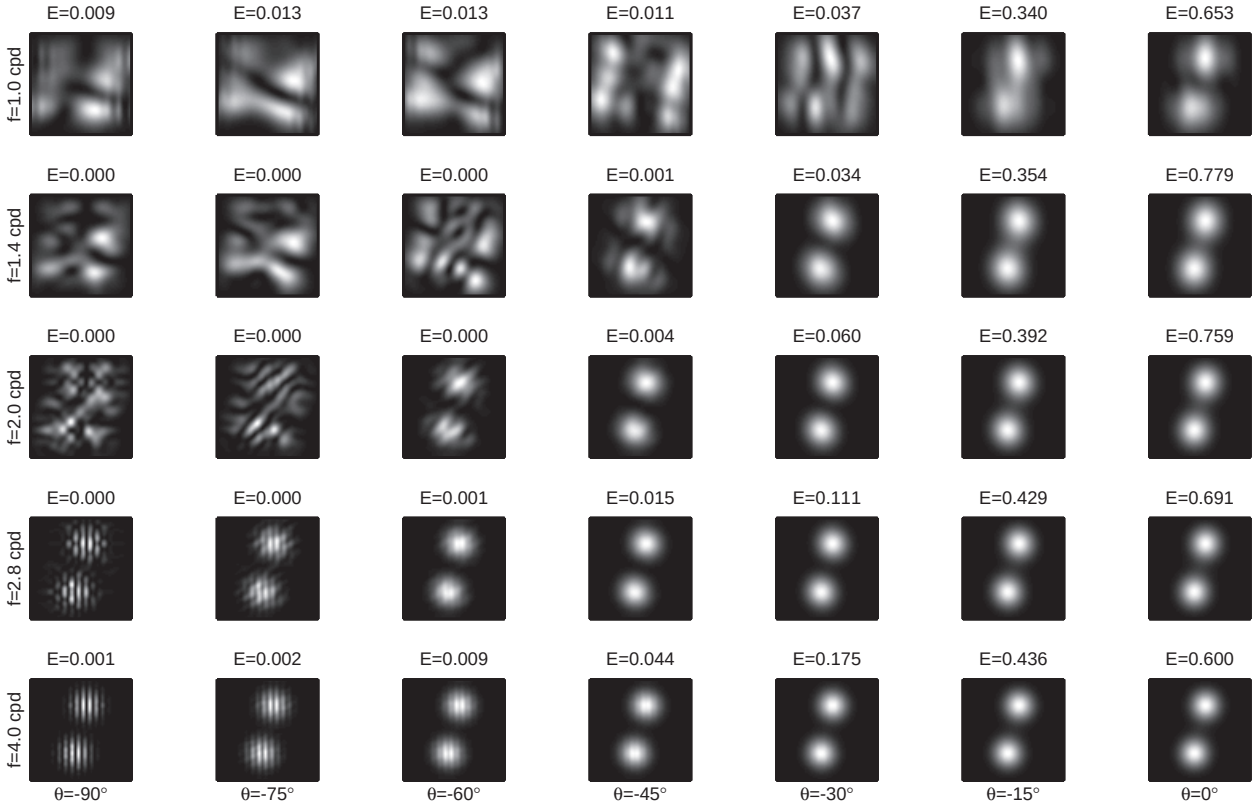

Figure B.11: Local energy maps of 5×7 channels from the V4 representation layer with the kernels shown in Figure B.10 (each row corresponds to 1 of 5 preferred spatial frequencies; each column corresponds to 1 of 7 preferred orientations), in response to a two-grating Vernier stimulus of offset 19 arcmin. The title above each map indicates the normalized and pooled (over space) energy of the unit, prior to the compressing nonlinearity. $\theta = 0°$corresponds to vertical; clockwise is positive.

AppendixC. Changes to the AHRM output activation function

One problem with using the original activation function of the AHRM (Eq. 4) for our tasks is that the resulting psychometric function (the probability of correct response as a function of stimulus difficulty/intensity – in this case offset or orientation/contrast difference) is not a sigmoid *separately* in the positive range (e.g. for offsets from 0 to 19) and in the negative (from 0 to -19) range. As Weiss et al. (1993) show (see also Eq. 11 in Petrov et al., 2005), the probability of a correct response in the model if its output (before an activation nonlinearity is applied) is linear in stimulus offset o is given by

$$P_{corr}(o) = \Phi(\frac{ao}{\sigma}) \tag{C.1}$$

where a is the slope of the output (as a function of signed offset) and σ is the standard deviation of the total noise (it can be shown that the noise, which is a combination of representation and decision noise, is Gaussian to a good approximation; if representation noise is neglected, the total noise equals the Gaussian decision noise: $\sigma = \sigma_d$). First off, this equation is only suitable for

positive offsets (or the absolute value of negative offsets); feeding this equation a negative offset will give a probability of less than 0.5, which does not make sense in a binary ("signal"/"no signal") experiment (as the minimum, chance, performance is 50%). Furthermore, $P_{corr}(o)$, given by Eq. C.1, from the hardest (zero) to the easiest (most positive or most negative) offset, will not produce a sigmoid, but the upper half of a sigmoid, which is a convex function, similar to a saturating exponential. This shape is retained even if the nonlinearity of Eq. 4 is applied, because tanh is itself convex for positive inputs. This is in direct conflict with human psychophysical data, where invariably the psychometric function in a binary/2-IFC experiment is a sigmoid, ranging from 0.5 for zero offset to 1 for the minimum negative or maximum positive offset. For the model to correctly reproduce human psychometric data under our particular choice of mapping of offsets to outputs, both the positive and the negative parts of its output activation need to have a sigmoid shape.

A second problem, related to the first, is that decision noise is added to the bottom-up input $\mathbf{w} \cdot \mathbf{X}$ before the activation function is applied, i.e. when the input is still linear in stimulus precision (e.g. offset). In the beginning of training, the slope of this offset-output line is relatively small. The problem with this became apparent in practice when the details of the staircase procedure in the simulation were varied. In particular, we noticed that when the number of reversals in a staircase was reduced from 20 (as in Hung & Seitz, 2014, apart from the orientation-only experiment) to 10 (as in the orientation-only experiment of Hung & Seitz, 2014 as well as in the experiments of Wang et al., 2014, which we simulate) the thresholds obtained by the model increased. This was because in the 20-reversal staircases, the first 10 are excluded from analysis whereas in the 10-reversal ones, only the first 4 are excluded. The first few reversals in a staircase are typically excluded from analysis as they occur at stimulus levels that are higher than the true threshold. In the model, the initially small slope of the linear bottom-up input, coupled with the fact that the noise was applied to the input before the activation function, resulted in a large number of initial reversals being above threshold – larger than what is seen in human data. A direct demonstration that in human data this effect is less prominent is seen in the orientation-only experiment of Hung & Seitz (2014), where test sessions are conducted with 20-reversal staircases and training sessions with 10-reversal ones (to allow comparison with results from other labs): the threshold in the test session in multiple-staircase mode was higher than the threshold in the first training session. Furthermore, the thresholds obtained by Wang et al. (2012) and Wang et al. (2014) in Vernier tasks are no higher than those obtained by Hung & Seitz (2014). In the model, however, this relationship was reversed

because the 10-reversal training staircases had caused it to converge to a value higher than the true threshold.

AppendixD. Parameters for the fitting of the IRT data set

One parameter that also needed to be adjusted and that was not reported in Table 1 of Dosher et al. (2013), but can be deduced from Figure S3 of the Supporting Information of that study, is the initial weight vector direction (Table 1 of Dosher et al., 2013 only reports the scale of the weight vector that determines its magnitude). As for differences in parameters between the 3 groups, Dosher et al. (2013) have only changed the V1 representation noise std.dev for the "Switch O" group, which was set to 0.02, as opposed to 0.01 in the other two groups. We did not vary this parameter across groups but instead chose to vary the initial weights scaling factor. This is because the only difference between the three groups in the Training phase (where all groups are performing the exact same condition) is the initial performance, as can be seen from the difference in performance in the first two blocks between the "Switch O" and the other groups, in *both* the noise-free and noisy conditions. Asymptotically, the three groups show no significant differences in the Training phase (blocks 4-8). Changing the initial weights makes intuitive sense, as it seems more likely that the 3 groups come from the same population and that, during initial exposure to the practice trials, the initial weights were set differently across groups, based on task demands: the "Switch O" group knew that there would not be a switch in position, and thus it might have given initial weights to V1 and V4 in a different way than the other groups. Our fits were obtained by just changing the V1 initial weight scaling factor (from 0.228 to 0.198) while keeping the respective V4 parameter at the same value across groups. Thus in the DPMM, as in the original model, only one parameter had to be changed in the "Switch O" group[4].

[4]We did not change, or use, the "internal noise 1" σ_1 parameter, which does not feature in all AHRM publications. This very small (of the order of 10^{-7}) parameter is not necessary in the DPMM, which includes internal noise (referred to as "internal noise 2" $\sigma_2$in Dosher et al., 2013).

AppendixE. Representation and decision unit activities

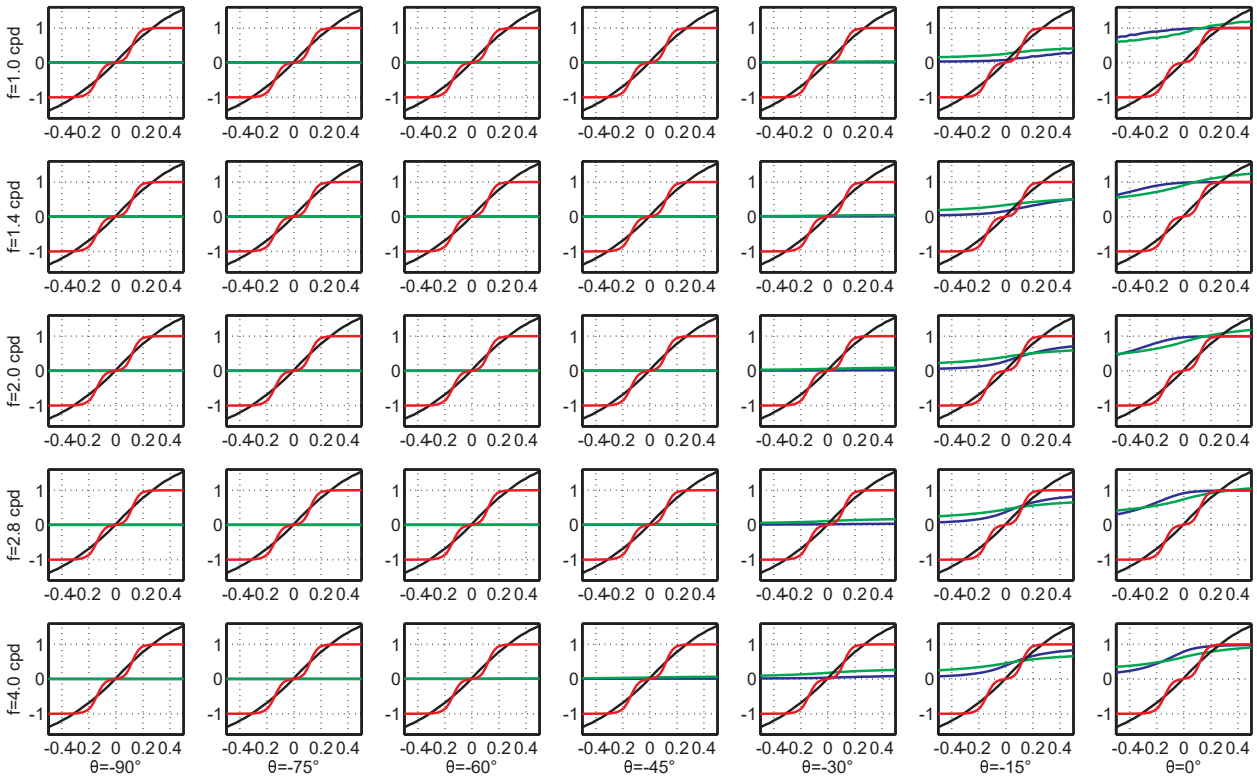

Figure E.12: Activities (ordinate) of the 5×7 representation units in V4 with the kernels shown in Figure B.10 and of the decision unit as a function of contrast difference ΔC (abscissa) between top and bottom grating in the contrast discrimination task. Green lines are the the raw outputs of the representation units (each having a particular orientation and spatial frequency preference); blue lines are the outputs passed through the representation activation function; the black line is the input to the decision unit ($\mathbf{w} \cdot \mathbf{X}$) that has been repeated in each subfigure for visual comparison; the red line is the decision unit output ($\mathbf{w} \cdot \mathbf{X}$ passed through the output activation function). Weights are post-training.

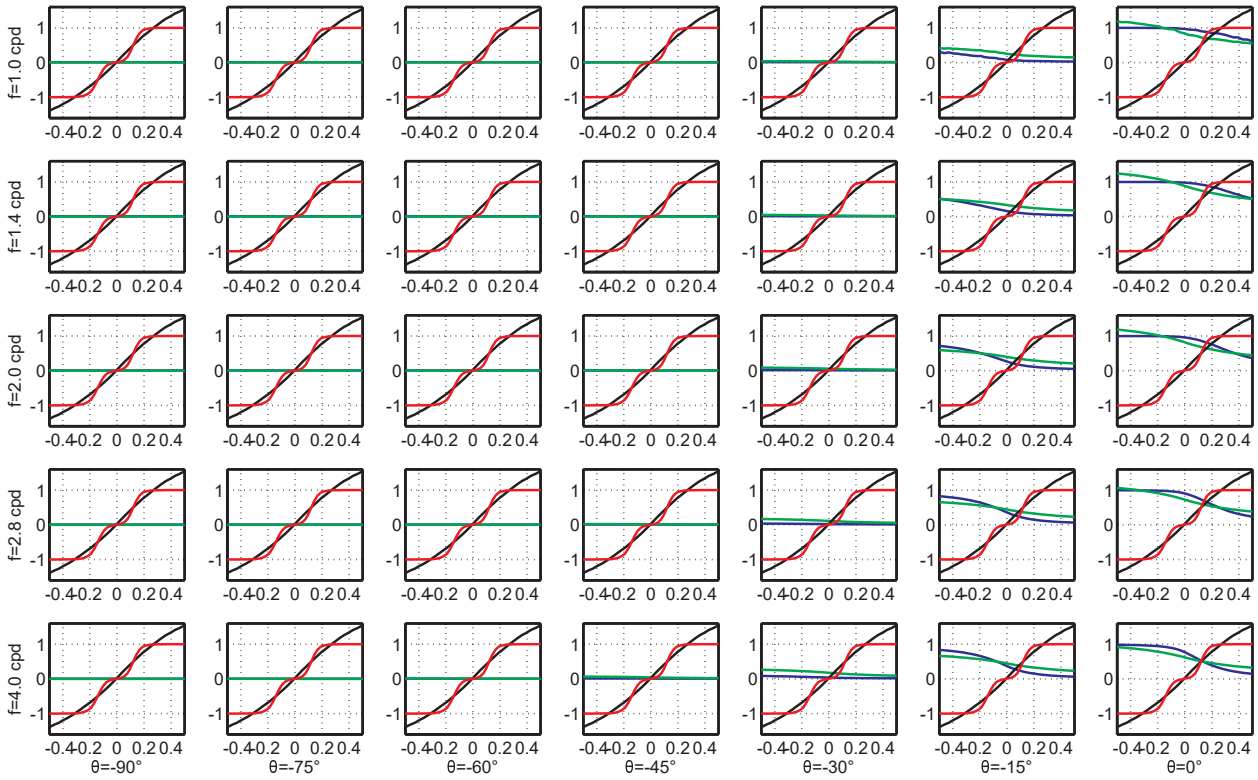

Figure E.13: Activities (ordinate) of the 5×7 representation units (each having a particular orientation and spatial frequency preference) in V4 with the "flipped" kernels (see 2.2.4) and of the decision unit as a function of contrast difference ΔC (abscissa) between top and bottom grating in the contrast discrimination task. Green lines are the the raw outputs of the representation units; blue lines are the outputs passed through the representation activation function; the black line is the input to the decision unit ($\mathbf{w} \cdot \mathbf{X}$) that has been repeated in each subfigure for visual comparison; the red line is the decision unit output ($\mathbf{w} \cdot \mathbf{X}$ passed through the output activation function). Weights are post-training.

F Figures of experimental data

Below are figures reproduced from the experimental studies of Hung & Seitz (2014) and Wang et al. (2014) that we simulated in the present work.

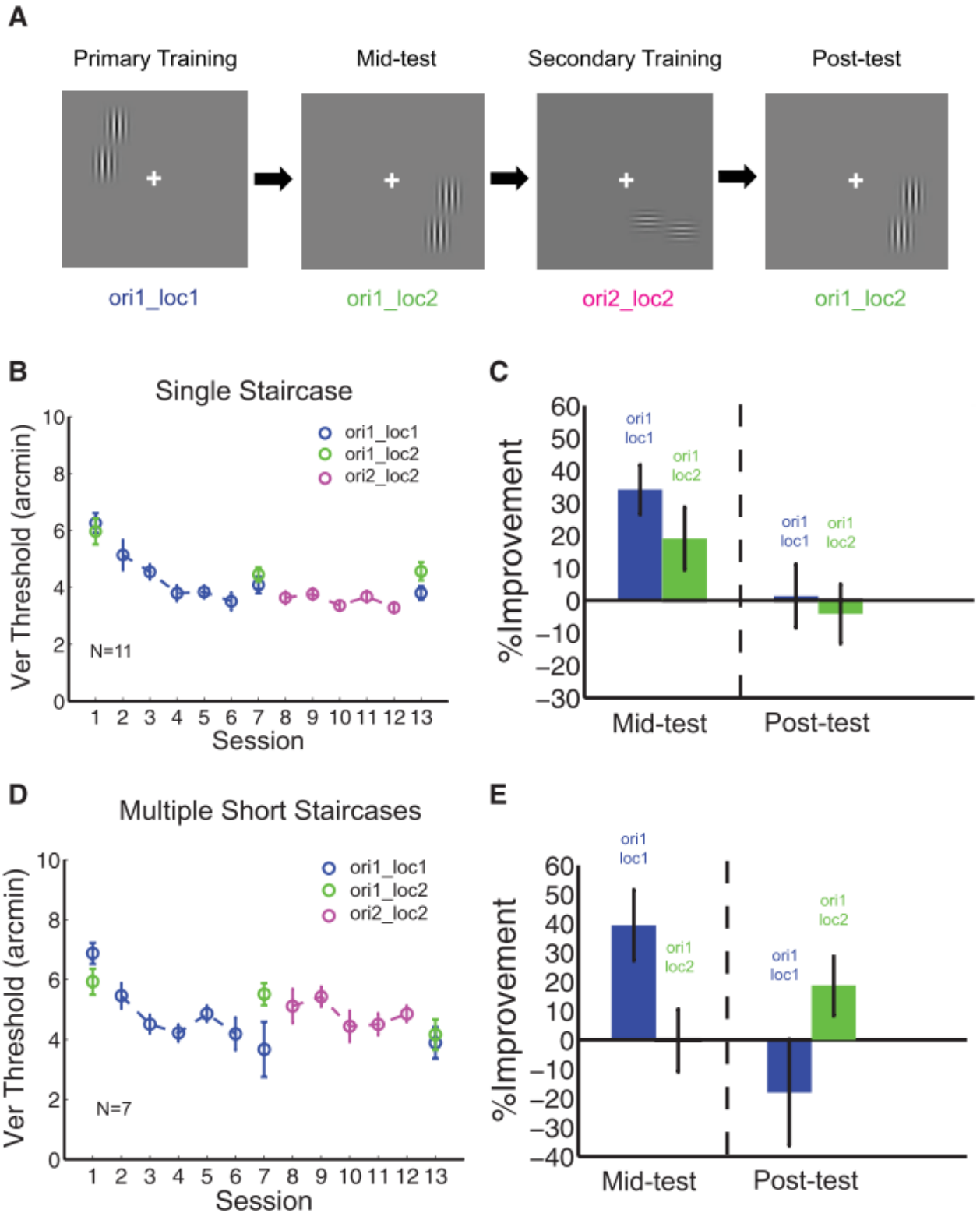

Figure E.14: Sequential double training data, adapted from Hung & Seitz (2014). A, schematic showing sequential double training. B, C, D, E correspond to A, B, C, D (respectively) of Figures 3 and 4.

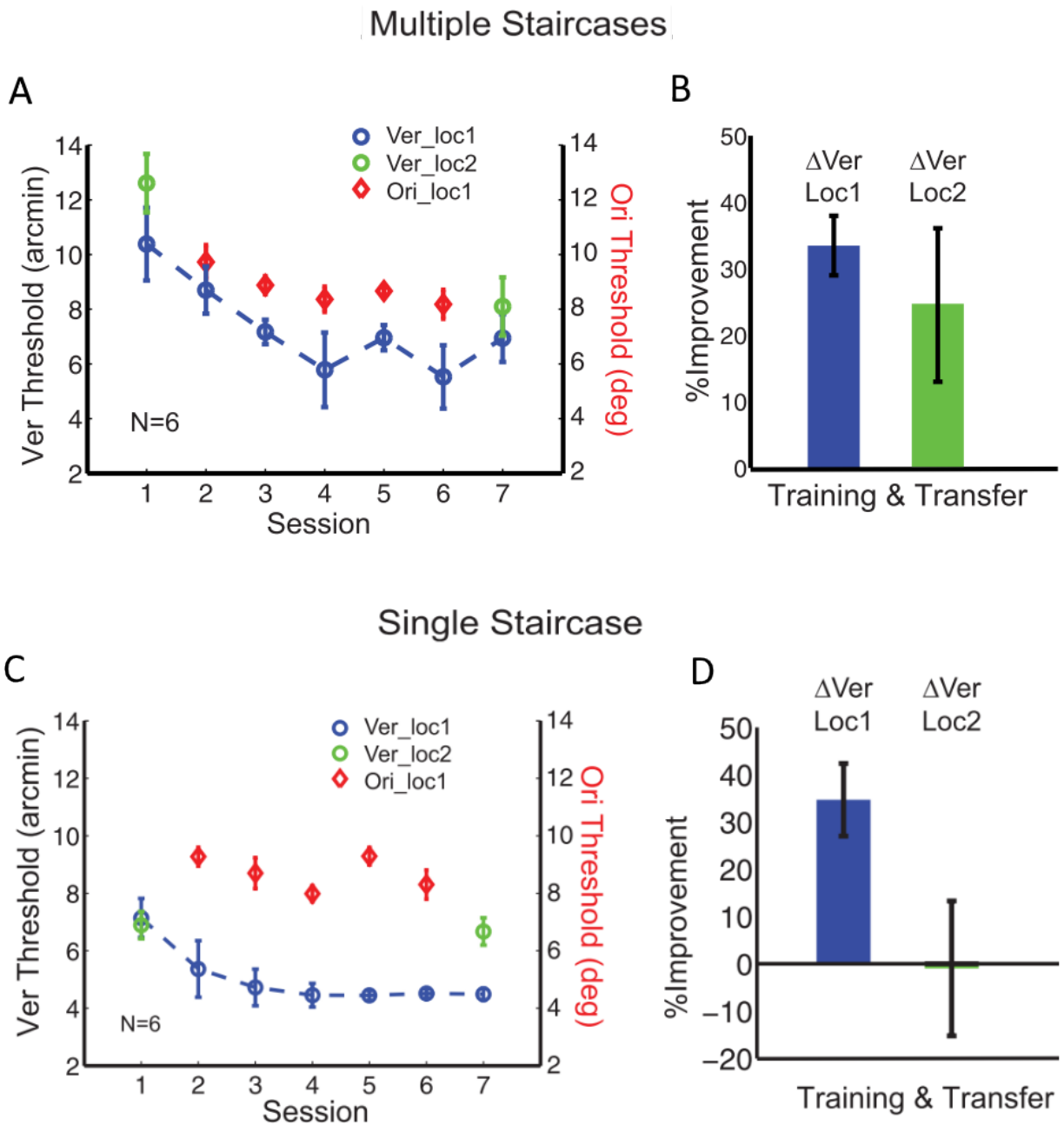

Figure E.15: Concurrent double training data, adapted from Figures 1 and 3 of Hung & Seitz (2014). Corresponds exactly to Figure 5.

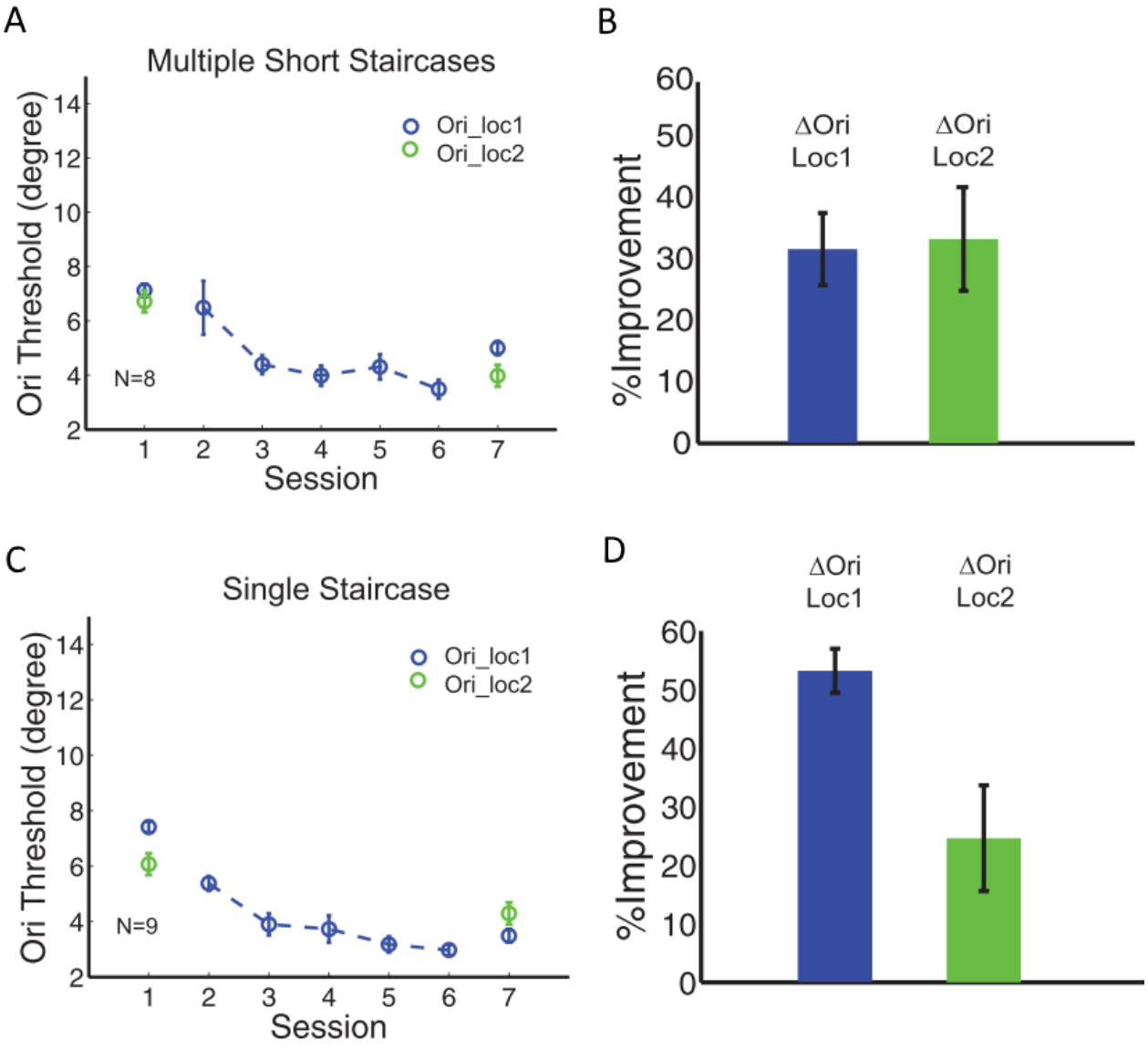

Figure E.16: Orientation training data, adapted from Figure 6 of Hung & Seitz (2014). Corresponds exactly to Figure 6.

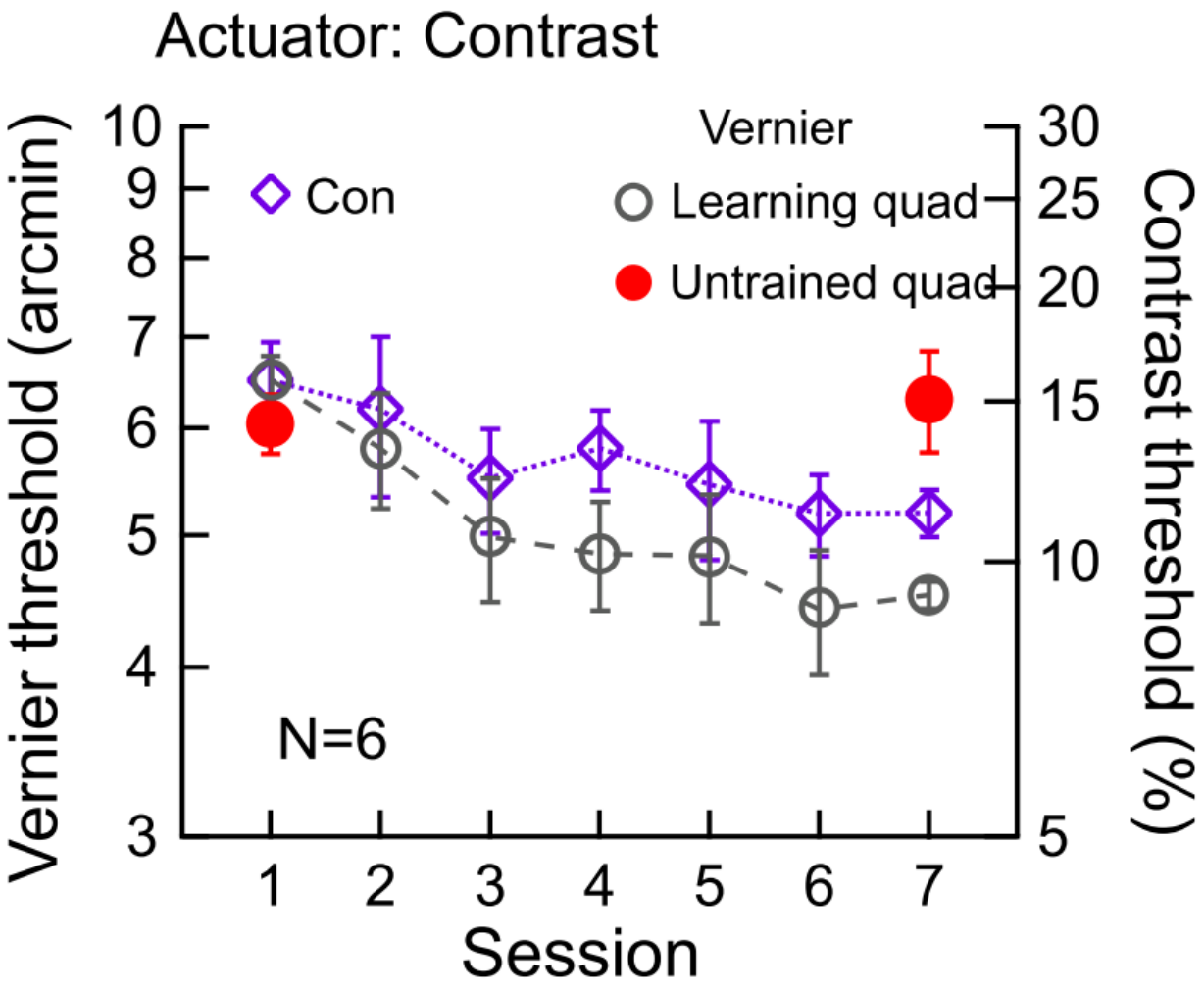

Figure E.17: Concurrent double training data, adapted from Figure 2c of Wang et al. (2014). Corresponds to Figure 6A.